

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 09:09:01 ; Search time 6808 Seconds
(without alignments)
5676.280 Million cell updates/sec

Title: US-09-617-099B-1
Perfect score: 8285
Sequence: 1 MSAPLGRGRPATPAASQP.....TRRASOSSLESSTGSPYSRS 1590

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09617099/runat_14112003_191546_23691/app_query.fasta_1.1735
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09617099@cgn_1_1_5271@runat_14112003_191546_23691 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1622	19.6	1346	11	AK0831172	AK0831172 Mus muscu
2	1522	18.4	1200	11	EC018057	EC018057 Homo sapi
3	1416	17.1	1073	13	BUS02896	BUS02896 AGENCOURT
4	1240	15.0	944	13	BUL140296	BUL140296 Mus muscu
5	1224.5	14.8	3687	11	AK032619	AK032619 Mus muscu
6	1157	14.0	698	14	CB527998	CB527998 UI-M-FY0-
7	1155	13.9	799	14	CB519497	CB519497 UI-M-GH0-
8	1149	13.9	686	10	BB428808	BB428808 BB428808
9	1085.5	13.1	915	13	B0413561	B0413561 602934853
10	1048.5	12.7	1004	14	CD301064	CD301064 AGENCOURT
11	1039	12.5	1718	11	AK043152	AK043152 Mus muscu
12	1037	12.5	697	13	BU741736	BU741736 UI-E-E01-
13	1006.5	12.1	1838	11	BC016274	BC016274 Mus muscu
14	1006.5	12.1	3080	11	BC037619	BC037619 Mus muscu
15	999	12.1	663	14	CA346437	CA346437 677333 NC
16	972.5	11.7	934	13	BUS030896	BUS030896 AGENCOURT
17	940.5	11.4	1684	11	BC024828	BC024828 Mus muscu
18	934.5	11.3	751	14	CB244903	CB244903 UI-M-FY0-
19	928.5	11.2	747	12	BI732490	BI732490 603353484
20	926	11.2	609	12	BM072612	BM072612 fv0907.Y
21	924.5	11.2	877	12	BI872362	BI872362 603396602
22	921.5	11.1	740	10	EG04784	EG04784 602420976
23	911	11.0	697	14	CD237320	CD237320 FNPJED04
24	908	11.0	603	2	HBM067824	Bx478187 Homo sapi
25	901.5	10.9	740	12	BI730821	BI730821 603351819
26	868	10.5	540	13	BUS38952	BUS38952 603513503
27	868	10.5	657	14	BY722517	BY722517 BY722517
28	863	10.4	762	10	EG675755	EG675755 602621981
29	855.5	10.4	1001	14	CA966376	CA966376 CCLX05a16
30	845	10.2	577	13	BU211560	BU211560 604159027
31	842	10.2	567	13	BU728456	BU728456 UI-E-CL1-
32	828	10.0	731	14	CB524260	CB524260 UI-M-FY0-
33	823.5	9.9	552	12	BM035114	BM035114 fu38804.Y
34	823	9.9	937	9	AL519002	AL519002 AL519002
35	820.5	9.9	1071	13	EX345439	EX345439 EX345439
36	820	9.9	939	13	EX331506	EX331506 EX331506
37	816.5	9.9	964	13	BUS23891	BUS23891 AGENCOURT
38	806	9.7	590	9	AA476826	AA476826 zw95901.Y
39	803	9.7	537	13	EX104117	EX104117 BX104117
40	787	9.5	542	12	BI844762	BI844762 fg17402.Y
41	784	9.5	474	9	AA772858	AA772858 ab70a10.X
42	782.5	9.4	754	10	EG207226	EG207226 RST26694
43	781	9.4	571	10	BE750318	BE750318 201646 MA
44	776	9.4	523	9	AA602733	AA602733 np17904.s
45	768	9.3	502	10	BE501603	BE501603 ht60g10.X

ALIGNMENTS

RESULT 1
AK0831172
LOCUS
DEFINITION
Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone:C630023J21 product:Rab3 interacting protein
1, full insert sequence.

AK0831172 1346 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone:C630023J21 product:Rab3 interacting protein
1, full insert sequence.

ACCESSION AK0831172
VERSION AK0831172.1 GI:26350308
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	1	Carninci, P. and Hayashizaki, Y.	Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
AUTHORS			
JOURNAL		High-efficiency full-length cDNA cloning	
MEDLINE		Meth. Enzymol. 303, 19-44 (1999)	
PUBMED		99279253	
PUBMED		10349636	
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
AUTHORS			Please visit our web site for further details.
TITLE		Normalisation and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	URL: http://genome.gsc.riken.go.jp/
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)	URL: http://fantom.gsc.riken.go.jp/
MEDLINE		20493374	
PUBMED		11042159	
PUBMED		11076861	
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	Location/Qualifiers
AUTHORS			1. .l346
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	/organism="Mus musculus"
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)	/mol_type="mRNA"
MEDLINE		20530913	/strain="CS7BL/6J"
PUBMED		11076861	/db_xref="FANTOM DB:C630023J21"
PUBMED		11076861	/db_xref="taxon:10090"
REFERENCE	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Harada, A., Fukunishi, Y., Komoto, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Tomita, M., Quackenbush, J., Schriml, L.M., Staehli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boifelli, D., Bojunga, N., Carninci, P., de Bonaldi, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seiya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boxis, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.	/clone_lib="RIKEN full-length enriched mouse cDNA library"
AUTHORS			/dev_stage="adult"
TITLE		Functional annotation of a full-length mouse cDNA collection	290. .>1345
JOURNAL		Nature 409 (6821), 685-690 (2001)	/note="unnamed protein product; Rab3 interacting protein 1 (MGD MG1:2152972, GB NM_053271, evidence: BLASTN, 100%, match=893)"
MEDLINE		21085660	/codon_start=1
PUBMED		11217851	/protein_id="BAC38794.1"
PUBMED		11217851	/db_xref="GI:26350309"
REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	/translation="MQFTLRQVNSVLSHPHFVSSPNTLONELFGOTLNNARKRS PSVSDQNRVEQSEREDSYVSDCTMPERSDSDYADRSQRSPQYEEPCHLNYS DNRGRHSHKEYIYDDSDVSRDSEYERQREEFQARYSDFNARIPFKQPTFEQ DNRHGVSRKHRRHSDVLSANAEISLRLMDRPSRVSERRAAMENQRSYK TNERTEAGQSSDPQTSNHSPTPRSPILDRPDMRRADSLRKHHLDPSSVRK TKRRTMETLNRDLSLSSQSSQSVRPPPPPHKSKGKVRQVLSLSSEELASTPEYT SCDDVELESSEVSEKGS"
AUTHORS			BASE COUNT 443 a 275 c 334 g 294 t
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	ORIGIN
JOURNAL		Nature 420, 563-573 (2002)	Alignment Scores:
MEDLINE		11217851	Pred. No.: 4,05e-103 Length: 1346
PUBMED		11217851	Score: 1622.00 Matches: 312
PUBMED		11217851	Percent Similarity: 100.00% Conservative: 0
REFERENCE			Best Local Similarity: 100.00% Mismatches: 0
AUTHORS			Query Match: 19.58% Indels: 0
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	DB: 11 Gaps: 0
JOURNAL		Nature 420, 563-573 (2002)	US-09-617-099B-1 (1-1590) x AK083172 (1-1346)
MEDLINE		11217851	QY 273 ArgLysArgSerProSerValSerArgAspGlnAsnArgArgTyrGluGlnSerGluGlu 292
PUBMED		11217851	DB 410 AGAAAACGAAGTCATCATGTGTCAGGGATCAAAATCGAAGATACGAGCAAGTGAAGAA 469
PUBMED		11217851	QY 293 ArgGluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAsp 312
REFERENCE			DB 470 AGAGAGGACTACTCAGATATGTTCTTCAGATGTGTACATGCCAAGATCTCTTCGGAT 529
AUTHORS			QY 313 TyrAlaAspArgArgSerGlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsn 332
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	DB 530 TATGCTGATAGCATCTCAGGCTGAGCTCAATTTTATGAAGACCTGGTCTTAAT 589
JOURNAL		Nature 420, 563-573 (2002)	QY 333 TyrArgAspSerAsnArgArgGlyHisArgHisSerLysGluTyrIleValAspAspGlu 352
MEDLINE		11217851	DB 590 TACAGGATCTTACACGAGAGGCGCATAGACATTCAAAGAGTATATTGTCGATGAA 649
PUBMED		11217851	QY 353 AspValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnAlaArg 372
PUBMED		11217851	DB 650 GATGTGGAGAGCAGAGATGAATATGAAACAAAGAGAGAGAGGAATACCAGGACGC 709
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
MEDLINE		11217851	
PUBMED		11217851	
PUBMED		11217851	
REFERENCE			
AUTHORS			
TITLE			

QY 373 TyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGln 392
 Db 710 TACAGAGTGTATCCAAATCTCGCCGGTATCCCGTAAAGCCACCAACCTACGAGAACAA 769

QY 393 MetArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSer 412
 Db 770 ATGGCATCCAGCTGAGGTCTCAGGCGACACATAGAGAGGACACAGTGAITTTCT 829

QY 413 LeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSer 432
 Db 830 TTGGCAACCGCTGAACTAGAAGATTCAGAGTTCCTCTGCTAAGATGATAGACCATCA 899

QY 433 ArgGlnArgSerValSerGluArgArgAlaMetGluAsnGlnArgSerTyrSerMet 452
 Db 890 AGGCAAGATCTGTATCTGACGCTAGAGTCAATGGAACCAACCATCGTATTCATG 949

QY 453 GluArgThrArgGluAlaGlnGlyLysSerSerTyrProGlnArgThrSerAsnHisSer 472
 Db 950 GAAAGAACTCGAGAGGCTCAGGACAAAGTTCCTATCCACAAAGCACTCAATCATAGT 1009

QY 473 ProProThrProArgSerProIleProLeuAspArgProAspMetArgAlaAsp 492
 Db 1010 COTCCACCCCTCGCGGAGCCCTATACCGCTTGATAGACCAACATGAGCGCGCTGAC 1069

QY 493 SerLeuArgLysGlnHisLeuAspProSerSerAlaValArgLysThrLysArgGlu 512
 Db 1070 TCCCTACGAAACAGCACCATTTAGATCCAGCTCTGCTGTGAGGAAACAGAGCGAGAA 1129

QY 513 LysMetGluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValArg 532
 Db 1130 AAAATGGAACCATGTTAAGGAATGATTCCTTTGAGTTTCAGACCACTCGAGTCAAGT 1189

QY 533 ProProProProArgProHisLysSerLysLysGlyLysMetArgGlnValSerLeu 552
 Db 1190 CGGCCCCCAACAGGCTCTATAATCCAAAGAAAGAGGATAAATGCCCGAGTTTCACTG 1249

QY 553 SerSerSerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspAspValGlu 572
 Db 1250 AGCAGCTCGAGAGGAGGAGCTGCATCCACACTGAGTATACAACTGTGATGATGAG 1309

QY 573 LeuGluSerGluSerValSerGluLysGlyAspSer 584
 Db 1310 CTGGAAGCGAGTGTGAGTGAGAAAGGGGACAGT 1345

RESULT 2
 BC018057
 LOCUS
 DEFINITION Homo sapiens, Similar to KIAA0751 gene product, clone
 IMAGE:4792768, mRNA.
 ACCESSION BC018057
 VERSION BC018057.1 GI:17390117
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1200)
 Strausberg, R.
 Direct Submission
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapps-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 32 Row: b Column: 12
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7662285
 This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

1..1200
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4792768"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DHIOB"
 /note="Vector: pBluescript"
 /note="Vector: pBluescript"

BASE COUNT 431 a 220 c 283 g 266 t
 ORIGIN

Alignment Scores:

Pred. No.: 3 38e-96 Length: 1200
 Score: 1522.00 Matches: 288
 Percent Similarity: 95.60% Conservative: 16
 Best Local Similarity: 90.57% Mismatches: 14
 Query Match: 18.37% Indels: 0
 DB: 11 Gaps: 0

US-09-617-099b-1 (1-1590) x BC018057 (1-1200)

QY 273 ArgLysArgSerProSerValSerArgAspGlnAsnArgArgTyrGluGlnSerGluGlu 292
 Db 242 AGGAAAGAAAGCCATCTGTCTCCAGATCAGATAGATAGATAGATAGATAGATAGATAG 301

QY 293 ArgGluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAsp 312
 Db 302 AGAGAGGAATATTCAGTATGCTACTTCGGATACCGCAATGCGCTAGATCTCCATCAGAT 361

QY 313 TyrAlaAspArgArgSerGlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsn 332
 Db 362 TATGCTGATAGCGGATTTCAACATGAACTTCAGTTTATGAAGACTCTGATCATTTAAGT 421

QY 333 TyrArgAspSerAsnArgArgGlyHisArgHisSerLysGluTyrIleValAspAspGlu 352
 Db 422 TATAGGAGCTCCCAACAGGAGAGTCTAGGCAATCCAAAGAAATATATTGTAGATCATGAG 481

QY 353 AspValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnAlaArg 372
 Db 482 GATGTGGAAAGCAGAGATGATACGAAAGGCAAGGAGAGAGAGAGTACCATGTCAGC 541

QY 373 TyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGln 392
 Db 542 TACCGAAGTGTATCCGAATTTGCGCGGTATTCAGTAAAGCCACAACTATGAAGAACAA 601

QY 393 MetArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSer 412
 Db 602 ATGCGGATCCATGCTGGAAGTGTCCGACGCGATGAGAGAGGATAGTGTATGTTCT 661

QY 413 LeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSer 432
 Db 662 TTGGCAACCGCTGATCTCGAAGATTCAGGATTTCTATGCTAAGATGATGATGATGAT 721

QY 433 ArgGlnArgSerValSerGluArgArgAlaMetGluAsnGlnArgSerTyrSerMet 452
 Db 722 AGGCAAGATCTATATCAGACGCTAGAGTGCCTGCAAGATGAGAGAGGATAGTGTATG 781

QY 453 GluArgThrArgGluAlaGlnGlyLysSerSerTyrProGlnArgThrSerAsnHisSer 472
 Db 782 GAAAGAACTCGAGAGGCTCAGGGACCAAGTCTTTATGCAACAAAGGACCAACCATAGT 841

473 ProProThrProArgSerProileProLeuAspArgProAspMetArgAlaAsp 492
 842 CCTCTACCCAGGAGGAGTCCACTACCATAGATAGACAGGAGTGTGAGTGC 901
 493 SerLeuArgGluHisLeuAspProSerSerAlaValArgLysThrLysArgGlu 512
 902 TCACACGGAACACACACACCTAGATCTCTAGCTCTCTGTGAAGAAAAACAAACGGGAA 961
 513 LysMetGluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerValArg 532
 962 AAAATGGAACAAATGTAAGGAATGATTTCTCTAGTTCCAGACAGTCCAGTGTGAGA 1021
 533 ProProProProArgProHisLysSerLysGlyGlyLysMetArgGlnValSerLeu 552
 1022 CCTCCACCAACCAAGCCATATAATCAAGAAAGCGGTAAATGCCCGCAGATTTCTGTTG 1081
 553 SerSerSerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspValGlu 572
 1082 AGCAGTTACAGAGGAGGAAATGGCTTCCACGCTGAATATACAACTGTGTGATGTTGAG 1141
 573 LeuGluSerGluSerValSerGluLysGlyAspSerGlnLysGlyLysArgLys 590
 1142 ATTGAAGTGCAGAGTGTAAAGTGAAGAAAGGGGACAGTCAAGAGGAAAAA 1195

RESULT 3

BU502896 1073 bp mRNA linear EST 12-SEP-2002
 AGENCOURT_8929724 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6489689
 5', mRNA sequence.
 BU502896
 BU502896.1 GI:22808299
 EST.
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1073)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Sequencing by: The I.M.A.G.E. Consortium (LIML)
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL4038 row: e column: 18
 High quality sequence stop: 571.

FEATURES

source
 1..1073
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6489689"
 /tissue_type="retina"
 /lab_host="PH108 (phage-resistant)"
 /clone_lib="NIH_MGC_94"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 340 a 238 c 277 g 217 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 7,44e-89 Length: 1073
 Score: 1416.00 Matches: 320
 Percent Similarity: 75.81% Conservative: 9
 Best Local Similarity: 73.73% Mismatches: 19

Query Match: 17.09% Indels: 86
 DB: 13 Gaps: 4
 US-09-617-099b-1 (1-1590) x BU502896 (1-1073)

354 ValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnAlaArgTyr 373
 DB 1 GTGAGAGCAGAGATGAATATGAAGACAAAGAGAGAGAGAGAGATACCGGCACGCTAC 60
 374 ArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMet 393
 DB 61 AGAAGTGATCCAAATCTGCCCGGTATCCCGTAAAGCCACCAACCTACGAGAACAAATG 120
 394 ArgLleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSerLeu 413
 DB 121 CGCATCCACGCTGAGGTGTCCAGGCGCACCATGAGAGAGAGAGAGAGATGTTCTTTG 180
 414 AlaAsnAlaGluLeuGluAspSerArgLleSerLeuLeuArgMetAspArgProSerArg 433
 DB 181 GCAAACGCTCAACTAGAGATTCAGGATTTCTCTGCTAAGGATGATAGACATCAAGG 240
 434 GlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGlu 453
 DB 241 CAAAGATCTGTATCTGAACGTAGAGCTGCAATGAGAAACCAACGATCGTATTCAATGAA 300
 454 ArgThrArgGluAlaGlnGlnSerSerTyrProGlnArgThrSerAsnHisSerPro 473
 DB 301 AGACTCCAGAGGCTCAGGACCAAGTCTTATCCACAAAGGACCTCAATCATAGTCT 360
 474 ProThrProArgSerProLleProLeuAspArgProAspMetArgArgAlaAspSer 493
 DB 361 CCCACCCCTCGCGGAGCCCTATACGCTGTATAGACCATGAGACATGAGGCGCGCTGACTCC 420
 494 LeuArgLysGlnHisHisLeuAspProSerSerAlaValArgLysThrLysArgLysLys 513
 DB 421 CTACGGAACACAGCACCATTTAGATCCAGCTCTGCTGTGAGAAACGAAGCGAGAAAA 480
 514 MetGluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValArgPro 533
 DB 481 ATGAAACCACTGTTAAGGAATGATTTCTGAGTTCAGACCATCCGAGTCAGTGAAGGCG 540
 534 ProProProArgProHisLysSerLysGlyGlyLysMetArgGlnValSerLeuSer 553
 DB 541 CCCCAACCAAGGCTCATTAATCCAGAAAGAGGAGTAANAATGCGCCAGGTTTCACTGAGC 600
 554 SerSerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspValGluLys 573
 DB 601 AGCTCGAGGAGGAGGAGTGGCATCCACCTGAGTATACAGCTGTGATGATGTGAGCT 660
 573 uGluSerGluSerValSerGluLysGlyAspSerGlnLysGlyLysArgLysThrSerG1 593
 DB 661 GGAAAGCGAGAGTGTGAGTGAAG----- 682
 593 uGlnGlyValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysArgMetTyrTy 613
 DB 682 ----- 682
 613 rGlyGlyHisSerLeuGluAspLeuGluTyrSerGluProGlnLleLysAspSerG1 633
 DB 682 ----- 682
 633 yValAspThrCysSerSerThrThrLeuAsnGluGluHisSerHisSerAspLys-Hisp 653
 DB 683 -----AAAGCACC 690
 653 roValThrTrpGlnProSerLysAspGlyAspArgLeuLleGlyArgLleLeuLeuAsnL 673
 DB 691 CTGTGACCTGCGACCCATCCAAAGATGGAGATCGCTTAATTTGGCGGTATTTATTAAATA 750
 673 ysArgLeuLysAsp-GlySerValProArgAspSerGlyAlaMetLeuGlyLysVal 692
 DB 751 AGCGTTTTAAAGATGGGAAGTGTACCTCGAGACTCAGGAGCAATGCTGGCGCTTAAGGCT 810

QY	693	Val - GlyGlyIysMetThrGluSer - GlyArgLeuCysAlaPheIleThrLysValLysL	712
Db	811	GTAAAGAAGAAAAGAACTGAATCCAGGCCACCTTTGGGCATTATTATTACCAAGTAAAA	870
QY	712	ysGlySerLeuAla - AspThrVal - GlyHisLeuArgProGlyAspGluValLeu - ---	729
Db	871	AGGAAATTTAGCTTGAAACTGTGAGACATCTTAACACAGGGGATGAATCTCTGGAAA	930
QY	730	- GluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPhe - GluGluValTrpAsnIleIle -	748
Db	931	GGAAATGGGAGGC - --TATTGCCAGGAGCCCTCTTTGGGGGAAGTTTACAACATTATTC	987
QY	749	LeuGluSerIysPro - GluProGlnValGluLeuValVal - SerArgProIleGlyAsp	767
Db	988	CTAAATTCACCAACTCGAACCCCAATTTGATCTGTGTGTTTCCAGGCCCATGTGGAAA	1047
QY	768	Ile - --ProArgIlePro 772	
Db	1048	TATTCTCTTAATATCCCC 1065	
RESULT 4			
LOCUS		BU140296	
DEFINITION		603134665F1 CSQCHL24 Gallus gallus CDNA clone CHEST116b17 5', mRNA	linear EST 25-NOV-2002
ACCESSION		BU140296	sequence.
VERSION		BU140296.1	GI:25355131
KEYWORDS		EST.	
SOURCE		Gallus gallus (chicken)	
ORGANISM		Gallus gallus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Archosauria; Aves; Neognathae; Galliformes; Phasianidae;	
		Phasianinae; Gallus.	
		1 (bases 1 to 944)	
		Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,	
		Fong, W. I., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.	
TITLE		A Comprehensive Collection of Chicken cDNAs	
JOURNAL		Curr. Biol. 12 (22), 1965-1969 (2002)	
MEDLINE		22335534	
PUBMED		12445392	
COMMENT		Contact: Simon Hubbard	
		Department of Biomolecular Sciences	
		University of Manchester Institute of Science and Technology (UMIST	
)	
		PO Box 88, Manchester, M60 1QD, UK	
		Tel: 01612008930	
		Fax: 01612360409	
		Email: Simon.Hubbard@umist.ac.uk.	
FEATURES		Location/Qualifiers	
source		1..944	
		/organism="Gallus gallus"	
		/mol_type="mRNA"	
		/strain="white Leghorn, Hisex"	
		/db_xref="taxon:9031"	
		/clone="CHEST116b17"	
		/dev_stage="16 day embryo"	
		/lab_host="DH10B"	
		/clone_lib="CSEQCHL24"	
		/note="Organ: brain; Vector: pBluescript II KS(+); Site 1:	
		EcoRI; Site2: NotI; Modification of pBluescript II KS(+)	
		[Stratagene] vector to accommodate cDNA produced with the	
		T-trimmed protocol (Construction of uni-directionally	
		cloned cDNA libraries from messenger RNA for improved 3'	
		end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387	
		,624). Cut pBluescript II KS(+) with NotI and EcoRI.	
		Ligate in double stranded adaptor containing BsgI and	
		BamHI sites [5'ggcgcgtcgagcccgatccgagaaaaag]	
		[5'aattcttttttcgagtcgcgggtgcacgc]	
BASE COUNT		321 a 212 c 236 g 175 t	
ORIGIN			
Alignment Scores:		1.23e-76	944
Pred. No.:			

AKO32619

LOCUS
DEFINITION Mus musculus adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:16430709C16 product:Rab3 interacting protein 1, full insert sequence.

ACCESSION AKO32619.1 GI:26328428

VERSION HTc; CAP trapper.

KEYWORDS HTc; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE Carninci P. and Hayashizaki Y.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE PUBMED 10349636

REFERENCE

AUTHORS Carninci P., Shibata Y.Y., Hayateau N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M. and Hayashizaki Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE PUBMED 20499374

REFERENCE

AUTHORS Carninci P., Itoh M., Aizawa K., Nagaoaka S., Saeki N., Carninci P., Konno H., Akiyama J., Ishii K., Kitsumai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasihagi K., Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara S., Watabiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A. and Hayashizaki Y.

TITLE Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE PUBMED 20530913

REFERENCE

AUTHORS Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyoasa H., Kondo S., Yamataka I., Saio T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikolaic D., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio F., Sakai K., Okido T., Furuno M., Aono H., Baldarelli G., Barsh G., Blake J.D., Bozellini D., Boujona N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Mazzionni L., Mashima J., Mazzaletti J., Mombaerts P., Nordone P., Ring B., Rigwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schonbach C., Seya T., Shibata Y., Storch K.P., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S. and Hayashizaki Y.

TITLE Functional annotation of a full-length mouse cdna collection

JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE PUBMED 21085660

REFERENCE

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length CDNA's

JOURNAL Nature 420, 563-573 (2002)
MEDLINE PUBMED 121217851

REFERENCE

AUTHORS Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukushima S., Furuno M., Hanagaki T., Hara A., Hashizume M., Hayashizaki K., Havatsuma N., Hiramoto K., Hirooka T., Hirozane T., Kawai J., Kawanabe T., Kohda M., Kuraku A., Miyajima T., Ohsato N., Okazaki Y., Osada S., Oyama K., Seki K., Shimizu T., Shirasawa T., Sogabe Y., Tagami M., Takaga A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M. and Hayashizaki Y.

DIRECT SUBMISSION Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Science Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gscc.riken.go.jp]
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216

COMMENT CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/

FEATURES source location/Qualifiers

CDS

polyA_signal polyA_site

BASE COUNT 1060 a 825 c 779 g 1023 t

ORIGIN

Alignment Scores:

Pred. No.: Length: 3687

Score: 1224.50 Matches: 266

Percent Similarity: 78.47% Conservative: 22

Best Local Similarity: 72.48% Mismatches: 47

Query Match: 14.78% Indels: 33

DB: Caps: 7

US-09-617-099B-1 (1-1590) x AK032619 (1-3687)

QY 1245 ArgThrSerSerAla---SerArgPheSe-----SerThr-SerTyrMetSerVal 1266
Db 9 CGGGACTTCACCCTGTGAAGCCGCCCTTTCTTGACACATTTCTCTCGCATCCCCCTCCCAGC 68
QY 1261 Gl nSerGl uArProAr GlyAnArGLyslleSe rValPh eThr-Se rLysMe tGlnAn 1281


```
/tissue type="Whole brain"
/dev stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP GH0"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction. Ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
BASE COUNT 255 a 176 c 189 g 176 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 8.7e-71 Length: 799
Score: 1155.00 Matches: 230
Percent Similarity: 97.87% Conservatative: 0
Best Local Similarity: 97.87% Mismatches: 5
Query Match: 13.94% Indels: 2
DB: 14 Gaps: 0
US-09-617-099b-1 (1-1590) x CBS19497 (1-799)
QY 273 ArglyserSerProSerValSerArgAspGlnAenArgArgTyrGluGlnSerGluGlu 292
DB 97 AGAAAAAGAGTCCATCATCATGTCAGGATCCAAATCGAAGTACGAGCAAGTGAAGAA 156
QY 293 ArgGluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAsp 312
DB 157 AGAGAGGACTACTCACAGTATGTTCTTCAGATGTCATCAATGCCAAGATCTCTTCGGAT 216
QY 313 TyrAlaAspArgSerGlnArgGluProGlnPheTyrGluGluProGlyHisLeuAen 332
DB 217 TATGCTGATAGACGATCTCAGCGTGAGCCTCAATTTTATGAAGAACCTGGTCATTAAAT 276
QY 333 TyrArgAspSerAsnArgGlyHisArgHisSerLyseGluTyrIleValAspAspGlu 352
DB 277 TACAGGGATTCTAACAGAGAGAGCCATAGACATTCCTCAAGAGTATATTGTGATGATGA 336
QY 353 AspValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnAlaArg 372
DB 337 GATGTGGAGAGCAGAGATGAATATGAAGACAAAGAGAGAGAGAGGAAATACCCAGGCAC 396
QY 373 TyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGln 392
DB 397 TACAGAGTGTATCCAAATCTGCCCGGTATCCCGTAAAGGCCAACCCCTACGAAAGAACAA 456
QY 393 MetArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSer 412
DB 457 ATGCGCATCCAGCTGAGGTGTCAGGGCAGCAGCATGAGAGAGGACAGTATGTTTCT 516
QY 413 LeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSer 432
DB 517 TTGGCAACGCTGAACATAGAGATTCAGAGATTCCTCTCTAAGGATGATAGACCATCA 576
QY 433 ArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrSerMet 452
DB 577 AGSCAAAGATCTGTATCTGAACGTAGAGCTGCAATCGAAAGAACCAACGATCTGATTCATG 636
QY 453 GluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnArgThrSerAsnHisSer 472
DB 637 GAAAGAACTCGAGAGGCTCAGGACANAGTTCTTATCCCANAGGACCTCANATCATAGT 696
QY 473 ProProThrProArgArgSerProIleProLeuAspArgProAspMetArgAlaAsp 492
```

```
121 GGTCAATTAAATACAGGGATCTTAAACAGAGAGGCCATAGACATTCCTCAAGAGTATATT 180
349 ValAspAspGluAspValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGlu 368
181 GTGATGATGATGAGATGTGGAGAGCAGAGATGAATATGAAGACAAAGAGAGAGAGGAA 240
369 TyrGlnAlaArgTyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnPro 388
241 TACCAGGACGCTACAGAGTGTATCCAAATCTGCCCGGTATCCCGTAAAGGCCAACACCC 300
389 TyrGluGluGlnMetArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHis 408
301 TACGAAGAACAAATCGCATCCAGCTGAGGTGTCAGGGCAGCAGCATGAGAGAGGAC 360
409 SerAspValSerLeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMet 428
361 AGTATGTTTCTTTGGCAACGCTGAACATAGAGATTCAGGATTTCTCTCTGTAAGGATG 420
429 AspArgProSerArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArg 448
421 GATAGACCATCAAGCAAGATCTGTATCTGAACGTAGAGCTGCAATCGAAACCAACGA 480
449 SerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnArgThr 468
481 TCGTATTCATGGAAGAACTCGAGAGGCTCAGGACAAAGTTCCTTATCCAAAGGACC 540
469 SerAsnHisSerProProThrPro-ArgArgSerProIleProLeuAspArgProAspWe 488
541 TCAATCATATGTCCTCCACCCCTCCGCGGAGCCCTATACCGCTTGATAGACACAT 600
488 LArgArgAlaAspSerLeuArgLyseGlnHisLeuAspProSerSerAlaValArgLy 508
601 GAGGCGCGTGACTCCCTACGGAACAGAC-CACITAGATCCAGCTCTGCTGTGAGGGA 659
508 sThrylsArgGluLysMetGluThrMetLeuArgAsn 520
660 AACGANGCAGAAAAAATGGAACCATGNTAGGAAT 696
```

```
RESULT 7
CBS19497
LOCUS
DEFINITION
  799 bp mRNA linear EST 28-MAR-2003
  UI-M-GH0-cef-f-20-0-UI.r1 NIH BMAP GH0 Mus musculus cDNA clone
  IMAGE: 6838773 5', mRNA sequence.
CBS19497
CBS19497.1 GI:29352852
EST.
Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 799)
  NIH-MGC http://mgi.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgabs-r@mail.nih.gov
  Tissue Procurement: Dr. Jim Lin, University of Iowa
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA sequencing by: Dr. M. Bento Soares, University of Iowa
  DNA sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/mousefl.html
  This clone was contributed by the Brain Molecular Anatomy Project
  (BMAP).
Seq primer: pYX-5.
Location/Qualifiers
  1..799
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6"
  /db_xref="taxon:10090"
  /clone="IMAGE: 6838773"
```

```
FEATURES
  source
```


AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs@mail.nih.gov
Tissue Procurement:
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM14590 row: e column: 18
High quality sequence start: 16
High quality sequence stop: 583.
Location/Qualifiers
1. .1004
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6959035"
/dev stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHG_XGC_Eyel"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection (XGC
) library."

FEATURES source

BASE COUNT 320 a 193 c 247 g 243 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 3.09e-63 Length: 1004
Score: 1048.50 Matches: 229
Percent Similarity: 86.42% Conservative: 32
Best Local Similarity: 75.83% Mismatches: 31
Query Match: 12.66% Indels: 13
DB: 14 Gaps: 3

US-09-617-099B-1 (1-1590) x CD301064 (1-1004)

QY 1295 SerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAla 1314
Db 1 AGTGTGAGTGGAGAAAA-TACCAGCTGGAAACACAAATGATGCTAGTCTGATACAGCC 59
QY 1315 ValGlyAlaLeuGlyThrSerGlyLysIleArgSerSerIleGlyAlaLysMetVal 1334
Db 60 GTTGGGACGATTGGACCTGTGGAAGAAAGGAGATCAAGTCTAAGTGCCAAAGTAGTT 119
QY 1335 AlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGly 1354
Db 120 GCTATAGTA-----TCACGAGAGAGCAGACATCCAGCTTACCCAAACAGATTCT 173
QY 1355 GlyGlyLysLysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGlu 1374
Db 174 GGCACACAGAAATTTAAAGATCTATTCAAGAGAGCAGACAGAAACAGCATGGCCAGCAGAA 233
QY 1375 MetArgAsnTrpMetThrArgGlnAlaSerArgGlnSerThrAspGlySerMetAsnSer 1394
Db 234 ATGAGAAGTCGGATGGTTCGGCACCCGAGCGCAGAGTCAACAGATGGGAGTATTAAACAGT 293
QY 1395 TyrSerSerGluGlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPhe 1414
Db 294 TACAGCTCAGAGGGCAATTTAAATTTTCCAGGGTGAGATTGGGGTTGATAGCAATTC 353
QY 1415 SerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThr 1434
Db 354 AGCGATTTCTGGATGGACTAGGACCCAGCACAACCTTA--GGCAGACAAACGTTAGCAACT 411

QY 1435 ProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGlu 1454
Db 412 CCAGCAATGGGGATATTGAGATGGATAGATAAAAGGGGAGTGTAGAGTGGAA 471
QY 1455 IleIleArgAlaArgGlyLeuValValLysProGlySerIleThrLeuProAlaProTyr 1474
Db 472 GTTATACGAGCAAGAGGTTTAAATACAAAGCCAGGCTCTAAATCAACTCCAGCTCTTAT 531
QY 1475 ValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLysLysLysThrLysValAla 1494
Db 532 GTAAAGGTATATCTTTTAGAAATGGAGCGTGTGTAGCCAGANAAGACAGAAATGCA 591
QY 1495 ArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGly 1514
Db 592 CGGAAAACCTGGACCTCTTTTACCACAGAACTTGGCTTTTGAAGAAAGTCCACATGG 651
QY 1515 ArgValLeuGlnIleIleValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMet 1534
Db 652 AAGGTCTCTCAGGTAAATGCTCTGGGAGAGCTATGGTCGAATGATGCATCACAATCTTATG 711
QY 1535 GlyValAlaGlnIleLeu-LeuAspGluLeuLeuSerAsnMetValIleGlyTrpPh 1554
Db 712 GGTGTTCACAGATCTTTTGGAGAACTCGATCTATCAAGTGTGTGATGGATGGTA 771
QY 1554 e-LysLeuPhePro-ProSerSerLeu-ValAspProThr---SerAlaProLeuThrAr 1572
Db 772 TTAATTTGTTTCGGCCCTTGGTCAATGGTGGATCCACGCTGACACCCCTTACCTCG 831
QY 1572 gArgAla-----SerGlnSerSerLeuGluSerSerThr-GlyProSerTyrSer 1598
Db 832 TAGGGG-CTTCCCAATTCTCTCTTTGAAAGTTTCTACAGGGTCTCTCCCTTTCA 886
RESULT 11
AK043152 1718 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
DEFINITION enriched library, clone:A730060M23 product:NIM2 homolog (Rattus
norvegicus), full insert sequence.
ACCESSION AK043152
VERSION AK043152.1 GI:26089509
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PubMed 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20493374
PubMed 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PubMed 11076861
REFERENCE 4

AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojohji, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Asburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, T., Pesole, G., Suzuki, K., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.P., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, C., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombærts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

TITLE
Functional annotation of a full-length mouse cDNA collection

JOURNAL
Nature 409 (6821), 685-690 (2001)

PUBMED
21085660

REFERENCE
11217851

AUTHORS
5

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
Nature 420, 563-573 (2002)

AUTHORS
6 (bases 1 to 1718)

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

AUTHORS

TITLE
Direct Submission

JOURNAL
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

source

Location/Qualifiers
1. .1718
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A730060M23"
/db_xref="taxon:10090"
/clone="A730060M23"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"

misc_feature

498..1423
/note="NIM2 homolog [Rattus norvegicus] (SPTB)[Q9JIR3, evidence: FASTA, 99.3%ID, 100%length, match=923] putative"

BASE COUNT 337 a 529 c 512 g 340 t
ORIGIN

Alignment Scores:

Pred. No.: 2,68e-62 Length: 1718
Score: 1039.00 Matches: 246
Percent Similarity: 61.80% Conservative: 63
Best Local Similarity: 49.20% Mismatches: 139
Query Match: 12.54% Indels: 52
DB: 11 Gaps: 12

US-09-617-099b-1 (1-1590) x AK043152 (1-1718)

Qy 1108 ArgThrThrThrArgSerArgSer-----SerGluArgProAspThrAsn 1122
Db 24 AGAACTGTGTGGCGACTGCGGAGCCAGCAGCGGCGGCGGAGCGAGCTGTCCGCC 83
Qy 1123 LeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAla--- 1141
Db 84 CGGTGGCGGGGATGCCCGCTGAGTCTCT---GTGCTGCCCTCCCGCGCGCTGCC 140
Qy 1142 LeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrProGly 1161
Db 141 CTACCCCTTCGAGTCCGCGC-----CCGCTCGGGCGCGCGCGCTACCGCGGGT 194
Qy 1162 Thr-----GlyArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGlu 1179
Db 195 CCGCGGGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 233
Qy 1180 ArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyLysGln 1199
Db 234 -----GCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 260
Qy 1200 ValAlaGlySerAspProArgLeuGluGlnAspTyHisSerLysTyArgSerGlyTyr 1219
Db 261 GCTCCAGGCGTGAGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 320
Qy 1220 AspProHisArgGlyAlaAspThrValSerThrLysSerSerSerAspValSerAsp 1239
Db 321 AAGCCTGCCCTTAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 380
Qy 1240 -ValSerAlaVal-----SerArgThrSerSerSerAlaSer 1250
Db 381 GCTTGTGCGATTCCTAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 440
Qy 1250 rArgPheSerSerThrSerTyMetSerValGlnSerGluArgProArgGlyAsnArgly 1270
Db 441 GCACCACTCAGACTCAGCAGCTTGAATTGCTTGAAGCCCTGCCATCTCTGGGGA- 495
Qy 1270 sIleSerValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLe 1290
Db 496 -----CCATGTTTAACTGGGAGCGCTGTCCAGCTCGGAGGGGCGGCGGCGG 548
Qy 1290 uThrLysSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGl 1310
Db 549 AGTAAGAGTTCACGATCAGCGGTGAGATCTGTGGATCCGAGCAAGCTGGAGGT- 603
Qy 1310 nSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgArgSerSer-IleG 1330
Db 604 -----GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 647
Qy 1330 lyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeu 1350
Db 648 GTGCCAAGATGTTAGCTATCTGTCGCTTGTGACCCAGTGGAGCAAAAGCACTTTCAG 707
Qy 1350 erGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerThrGluThrG 1370
Db 708 CCCAGCTGAAGGGGCCACCAAGAGCTTTCGAAGCAACATCCGAGGACGACGAGACTG 767
Qy 1370 lyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThrAspG 1390
Db 768 GCATTGCAGTGGAGATCGGAGCCCGGTGTAACCCCGGCGGCGGCGGCGGCGGCGG 827

Db	276	AGATCATCGTCTGGGGAGATTATGGCCGCATGGATCAAAATCTTTTATGGGAGTGCCCC	217
Qy	1538	lnlleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGlyTrpPheIleValLeuPheP	1558
Db	216	AGATACTTTTAGATGAACCTAGAGCTATCCAAATATGGTGATCGGATGGTTCAACTTTTCC	157
Qy	1558	roProSerSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerS	1578
Db	156	CACCTTCTCCCTAGTAGATCCAACTTGGCCCTCTGACAAAGAAGAGCTTCCCAATCAT	97
Qy	1578	erLeuGluSerSerThrGlyProSerTyrSerArgSer	1590
Db	96	CTCTGGAAAGTTCAACTGGACCTTCTACTCTCGTTCA	59

RESULT	13
BC016274	
LOCUS	linear
DEFINITION	Mus musculus, clone IMAGE:4527810, mRNA.
ACCESSION	BC016274
VERSION	BC016274.1 GI:16740819
KEYWORDS	HTC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria;	Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 1838)
AUTHORS	Straussberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (29-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov

Percent Similarity:	90.34%	Conservative:	20
Best Local Similarity:	81.93%	Mismatches:	22
Query Match:	12.15%	Indels:	1
DB:	11	Gaps:	1

US-09-617-099B-1 (1-1590) x BC016274 (1-1838)

QY	1353	GlUGlYGLyGlyLysLysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAla	1372
Db	105	GAGTCGGCGCACAGAAGTTAAAGACCACTTACAGGAGTAGCGAACCAGGAATGGCG	164
QY	1373	ValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThrAspGlySerMet	1392
Db	165	GCTGAATGCGGAAG--ATGGTGACACAGCCAGCCGGGAGTCCACGGATGCAGCATC	221
QY	1393	AsnSerTyrsSerGluGlyAsnLeuIlePheProGlyValargLeuAlaSerAspSer	1412
Db	222	AACAGTTATAGCTCAGAAGGAAACTTAATAATTCTCGGAGTTCGCGTGGGACCTGCAGT	281
QY	1413	GlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeu	1432
Db	282	CAGTTTCAGTGATTTCCTTGATGATTGGGGCCAGCCAGCTGTTCGGCCGCCAACCCCTC	341
QY	1433	AlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspTylsGlyGlnLeuGlu	1452
Db	342	GCCACCCCACCCATGGGCGATATCCAATCGGATGGAGGATAAGAGGGTTCAGTTGGAG	401
QY	1453	ValGluIleIleArgAlaArgGlyLeuValVallysProGlySerLysThrLeuProAla	1472
Db	402	GTTGAGGTTATCAGAGCCCGAGCGCTTACACAAAACCTGGCTCCAANTCTACACCTGCT	461
QY	1473	ProTyrVallysValTyrLeuLeuAspAshGlyValCysIleAlaLysLysLysThrLys	1492
Db	462	CCCTATGTGAAGTTATCTTTTGGAAAAATGGAGCCTGTATTGCCAAAAAGAACACAAGA	521
QY	1493	ValalaArgLysThrLeuGluProLeuTyrGlnGlnLeuSerPheGluGluSerPro	1512
Db	522	ATTACCGGAAAACTCTCGATCCCTTGATCAGCAGTCCCTGGTTTTTGTGAAGATGCCA	581
QY	1513	GlnGlyArgValLeuGlnIlelleValITpGlyAspTyrGlyArgMetAspHisLysSer	1533
Db	582	CAGGTAAGAGTCTTCAGTGATTGTCTGGGGTCACTATGGAAGATGACACCAATATGC	641
QY	1533	PheMetGlyValAlaGlnIleLeuLeuAspGluLeuLeuSerAsnMetValIlegly	1553
Db	642	TTTATGGGTGGCTCAAACTCTTGTGAAGAACTTGATCTGTCCAGCATGGTGAATGGA	701
QY	1553	TrpPheLysLeuPheProSerSerLeuValAspProThrSerAlaProLeuThrArg	1573
Db	702	TGGTATAAATTTGTTCCCTCCATCTCTCTGGTGGATCCCACTCACTCCCTGACCCGC	761
QY	1573	ArgAlaSerGlnSerSerLeuGluSerThrGlyProSerTyrsSerArgSer	1590
Db	762	CGGGCTTCCCAATCATCACTGGAAAGTTCTGTGGCCCTCCCTGTCATCCGGTCA	815

RESULT 14
BC037619
LOCUS
DEFINITION Mus musculus clone IMAGE:5400144. mRNA.
BC037619 3080 bp linear HFC 04-MAR-2000

	RESULT 14	
	BC037619	
	LOCUS	
	DEFINITION	
	ACCESSION	
	VERSION	
	KEYWORDS	
	SOURCE	
	ORGANISM	
	REFERENCE	
	AUTHORS	
	TITLE	
	JOURNAL	
	BASE COUNT	
	ORIGIN	
FEATURES		
source		
	1..1838	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/db_xref="taxon:10090"	
	/clone="IMAGE:4527810"	
	/tissue_type="Eye, retina, mouse strain C57BL\6"	
	/clone_lib="NIH MGC_94"	
	/lab_host="DH10B"	
	/note="Vector: pCMV-SPORT6"	
	508 a 408 c 416 g 506 t	

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Legaspi, R.,
 Madurow, J.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>
 Series: IRAX Plate: 81 Row: p Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, GenomeScan gene prediction
 This clone has the following problem: retained intron.

FEATURES

source
 Location/Qualifiers
 1..3080
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:5400144"
 /tissue_type="Eye, retina, mouse strain C57BL/6"
 /clone_lib="NIH MGC_94"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 BASE COUNT 915 a 648 c 632 g 885 t
 ORIGIN

Alignment Scores:

Pred. No.: 9,86e-60 Length: 3080
 Score: 1006.50 Matches: 195
 Percent Similarity: 90.34% Conservative: 20
 Best Local Similarity: 81.93% Mismatches: 22
 Query Match: 12.15% Indels: 1
 DB: 11 Gaps: 1

US-09-617-099B-1 (1-1590) x BC037619 (1-3080)

QY 1353 GlnGlyGlyLysLysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAla 1372
 Db 123 GAGTCCGGCCACAGAAGATTAAAGACCAATTCAGAGGAGTACGGAACGGGAATGGCG 182
 QY 1373 ValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThrAspGlySerMet 1392
 Db 183 GCTGAAATCGGGAAG---ATGGTGAGACAGCCCGCGGAGTCCAGGATGGCAGCATC 239
 QY 1393 AsnSerTyrSerSerGluGlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSer 1412
 Db 240 AACAGTTATAGCTCAGAGGAAGAACTTAATTTCTCGAGTTCGCGGGGACCTGACAGT 299
 QY 1413 GlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeu 1432
 Db 300 CAGTTTCAGTGATTTCCITGATGGATTGGGGCCAGCCAGCTTGTGCGCCGCAACCCCTC 359
 QY 1433 AlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysGlyGlnLeuGlu 1452
 Db 360 GCCACCCAGCATCGGCGGATATCCAAATCGGATGGAGGATAAGAGGGTTCAGTTGGAG 419
 QY 1453 ValGluIleIleArgAlaArgGlyLeuValValLysProGlySerLysThrLeuProAla 1472
 Db 420 GTTGAGGTTATCAGAGCCCGGAGCCTTACACAAACCTGGCTCCAAATCTACACCTGCT 479

QY 1473 ProTyrValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLysLysLysThrLys 1492
 Db 480 CCTATGTGAAGTTTAACTTTTGGAAATGGAGCTGTATTGCCAAAAGAGACAGA 539
 QY 1493 ValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerPro 1512
 Db 540 ATTGCACGGAAACTCTCGATCCCTTGTATCAGCAGTCCTCGTGTGTTTATGATAAAGTCCA 599
 QY 1513 GlnGlyArgValLeuGlnIleIleValTrpGlyAspTyrGlyArgMetAspHisLysSer 1532
 Db 600 CAGGGTAAAGTTCTTCAGGTGATTGTCTGGGTGACTATGGAAGAAATGACCAACAATGC 659
 QY 1533 PheMetGlyValAlaGlnIleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGly 1552
 Db 660 TTATGGGTGTGGCTCAATCTTGTGGAAAGAACTTGATCTGTCCAGCATGTTGATTGA 719
 QY 1553 TrpPheLysLeuPheProSerSerLeuValAspProThrSerAlaProLeuThrArg 1572
 Db 720 TGTATAAATTTGTCCTCATCTCTCTGTTGGATCCACACTCACTCCCTGACCCGC 779
 QY 1573 ArgAlaSerGlnSerSerLeuGluSerThrGlyProSerTyrSerArgSer 1590
 Db 780 CGGGCTTCCCAATCATCACTGGAAGTTCTGTGGGCTCTCCCTGCATCCGGTCA 833

RESULT 15

CA346497 663 bp mRNA linear EST 05-NOV-2002
 LOCUS 677393 NCCOWA 1RT Oncorhynchus mykiss cDNA clone lRT101D11_B_B06
 DEFINITION 5', mRNA sequence.
 ACCESSION CA346497
 VERSION CA346497.1 GI:24591668
 KEYWORDS EST.
 SOURCE Oncorhynchus mykiss (rainbow trout)
 ORGANISM Oncorhynchus mykiss
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 REFERENCE 1 (bases 1 to 663)
 AUTHORS Rexroad, C.E. and Keele, J.W.
 TITLE Sequence analysis of a rainbow trout normalized cDNA library
 JOURNAL Unpublished
 COMMENT Contact: Rexroad CE
 USDA, ARS, National Center for Cool and Cold Water Aquaculture
 11876 Leetown Road, Kearneysville, WV 25430, USA
 Tel: 304 724 8340 x2129
 Fax: 304 725 0351
 Email: crexroad@nccowa.ars.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified by
 cross_match v0.990329.

FEATURES

Seq primer: ATTAGGTGACACTATAG.
 Location/Qualifiers
 source
 1..663
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="lRT101D11_B_B06"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="NCCOWA 1RT"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from brain, gill, liver,
 spleen, muscle, and kidney."
 BASE COUNT 170 a 136 c 173 g 124 t
 ORIGIN

Alignment Scores:

Pred. No.: 5.41e-60 Length: 663
 Score: 999.00 Matches: 192
 Percent Similarity: 91.28% Conservative: 7
 Best Local Similarity: 88.07% Mismatches: 3
 Query Match: 12.06% Indels: 16

DB:	14	Gaps:	1
US-09-617-099B-1 (1-1590) x CA346497 (1-663)			
Qy	753	ProGluProGlnValGluLeuValValSerArgProIleGlyAspIleProArgIlePro	772
Db	3	CCAGAACCCCGAGTGGTGGTCTCTAGAGCCCATAGGAGATATTCCTCGATACCA	62
Qy	773	AspSerThrHisAlaGlnLeuGluSerSerSerPheGluSerGlnLysMetAsp	792
Db	63	GACAGCACACATGCACAACATCGAATCAAGTTGAGTTCCTTTGAGTCTCAGAGATGAT	122
Qy	793	ArgProSerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGln	812
Db	123	CGCCCTCTATCTCTGTGTCACGTCTCCCATGAGCCCGGGATGCTGCGGAGCGCCCCCAG	182
Qy	813	PheLeuSerGlyGlnLeuSer-----	819
Db	183	TACCTGTGACGACAGCTCTCAAGCCAAAGCCTTAGTAGAACAACAGCGCCTTTGTCCCT	242
Qy	820	-----IleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeu	836
Db	243	AGGTCCAGGTCAAACTGTGTATGACAAAGTGGGCCATCAGCTAATAGTACCATCCTG	302
Qy	837	GlyAlaLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIle	856
Db	303	GGAGCCAAAGGACCTGCCTTCTAGGGAAGATGGCCGCCCGGAAACCTTACGTCAAAATC	362
Qy	857	TyrPheLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThr	876
Db	363	TACTTCTCCTGACAGAGCGGACAAAGCAAGAGAGAGAGACAAACCGTAAGAAATCC	422
Qy	877	LeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArg	896
Db	423	TTAGAGCCCAATGGAACAGACCTTATGTATTCTCCTGTGCACCGCGGGAGTTCGT	482
Qy	897	GluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGlu	916
Db	483	GAGCGCATGTTGGAGATCACCTGTGGGACCCAGGCCAGGGTCAGGGAGAGAAAGCGAG	542
Qy	917	PheLeuGlyGluIleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrp	936
Db	543	TTCTGGGAGAGATTAAGTAGAGCTGGAGACAGCCCTGCTGTGATGATGAGCTCACTGG	602
Qy	937	TyrLysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerPro	954
Db	603	TACAGCTGCAGACCCACAGCGTGTCTCCATGCCCTCCCGCGCCTCGGCC	656

Search completed: November 16, 2003, 15:02:58
Job time : 6846 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 09:19:06 ; Search time 1020 Seconds

(without alignments)

5094.554 Million cell updates/sec

Title: US-09-617-099B-1

Perfect score: 8285

Sequence: 1 MSAPLGRGRPAFTPAASQP.....TTRASQSSLESSTGTSYSRS 1590

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 433922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09617099/runat_14112003_191547_23738/app_query.fasta_1.1735
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCI=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09617099@cgn2_1_1708@runat_14112003_191547_23738
-NCFU=6 -ICFU=3 -NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	8285	100.0	4980	14	US-10-180-326-2 Sequence 2, Appli

2	2936.5	35.4	6638	12	US-10-220-891-40	Sequence 40, Appl
3	1042	12.6	7325	12	US-10-240-965-72	Sequence 72, Appl
4	374	4.5	9027	12	US-09-930-213-304	Sequence 304, App
5	368.5	4.4	24387	11	US-09-764-891-9844	Sequence 9844, Ap
6	332	3.8	9025608	14	US-10-156-761-1	Sequence 1, Appli
7	300.5	3.6	9169	12	US-09-814-353-20083	Sequence 20083, A
8	295.5	3.6	9025608	14	US-10-156-761-1	Sequence 1, Appli
9	295	3.6	9161	12	US-10-133-937-37	Sequence 37, Appl
10	295	3.6	11950	12	US-10-252-157-218	Sequence 218, App
11	288.5	3.5	52101	12	US-10-132-134-1	Sequence 1, Appli
12	284.5	3.4	4215	14	US-10-177-293-448	Sequence 448, App
13	280	3.4	11295	12	US-09-814-353-19153	Sequence 19153, A
14	279	3.4	12227	14	US-10-177-293-422	Sequence 422, App
15	270.5	3.3	5418	14	US-10-156-761-5959	Sequence 5959, Ap
16	270.5	3.3	7997	9	US-09-864-761-17726	Sequence 17726, A
17	268	3.2	9579	12	US-10-132-134-1	Sequence 11, Appl
18	265.5	3.2	10383	12	US-09-960-253-181	Sequence 181, App
19	265.5	3.2	10383	12	US-10-007-926A-56	Sequence 56, Appl
20	264.5	3.2	15872	10	US-09-861-289-1	Sequence 1, Appli
21	264.5	3.2	15872	10	US-09-860-846-1	Sequence 1, Appli
22	264.5	3.2	15872	11	US-09-988-384B-1	Sequence 1, Appli
23	264.5	3.2	15872	11	US-09-836-821-1	Sequence 1, Appli
24	264.5	3.2	15872	12	US-10-271-883-44	Sequence 44, Appl
25	264	3.2	77536	12	US-09-940-316B-1	Sequence 1, Appli
26	260.5	3.1	68750	13	US-10-014-717-1	Sequence 1, Appli
27	260	3.1	10174	14	US-10-171-311-82	Sequence 82, Appl
28	260	3.1	42999	10	US-09-799-462A-17	Sequence 17, Appl
29	260	3.1	42999	11	US-09-836-911A-17	Sequence 17, Appl
30	260	3.1	42999	12	US-09-738-630-73	Sequence 73, Appl
31	260	3.1	42999	13	US-10-125-767-17	Sequence 17, Appl
32	260	3.1	42999	14	US-10-151-081-17	Sequence 17, Appl
33	260	3.1	42999	14	US-10-287-313-17	Sequence 17, Appl
34	260	3.1	42999	14	US-10-219-694-17	Sequence 17, Appl
35	260	3.1	88421	10	US-09-976-059-1	Sequence 1, Appli
36	259	3.1	3651	12	US-10-311-406-3	Sequence 3, Appli
37	257	3.1	38734	11	US-09-373-658-30	Sequence 30, Appl
38	256	3.1	14835	12	US-10-240-965-113	Sequence 113, App
39	255.5	3.1	6457	10	US-09-880-107-3389	Sequence 3389, Ap
40	255.5	3.1	7497	10	US-09-960-253-175	Sequence 175, App
41	255.5	3.1	10056	14	US-10-156-761-411	Sequence 411, App
42	255.5	3.1	100000	14	US-10-156-761-15103	Sequence 15103, A
43	254.5	3.1	7058	12	US-10-287-218-38	Sequence 38, Appl
44	254.5	3.1	7792	13	US-10-044-090-359	Sequence 359, App
45	254	3.1	4974	14	US-10-156-761-2487	Sequence 2487, Ap

ALIGNMENTS

RESULT 1
US-10-180-326-2
; Sequence 2, Application US/10180326
; Publication No. US20030049661A1
; GENERAL INFORMATION:
; APPLICANT: Seino, Susumu
; APPLICANT: Shibusaki, Tadao
; APPLICANT: Ozaki, No. US20030049661A1uaki
; TITLE OF INVENTION: Protein Rim2
; FILE REFERENCE: P21573
; CURRENT APPLICATION NUMBER: US/10/180,326
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: JP 288372/99
; PRIOR FILING DATE: 1999-10-08<160> 5
; NUMBER OF SEQ ID NOS:
; SOFTWARE: Patent version 3.0
; SEQ ID NO 2
; LENGTH: 4980
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-180-326-2

Alignment Scores: 0 Length: 4980
Pred. No.: 8285.00 Matches: 1590
Score:

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-617-099B-1 (1-1590) x US-10-180-326-2 (1-4980)

Qy 1 MetSerAlaProLeuGluYrProArgGlyYrGProAlaProThrProAlaAlaSerGlnPro 20
Db 32 ATGTCGGCTCGCTCGGGCCCGGGCCCGCGCTCCACACCCCGCGCGCTCTCAACCT 91
Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIlelle 40
Db 92 CCTCGCAGCCGAGATGCCGACCTCAGCCACCTCAGAGAGAGAGAGAGAGAAATCATC 151
Qy 41 LeuAlaValMetAspArgGlnLysGluGluGluGluGluGluGluGluGluGluGlu 60
Db 152 CTGGCTGTCATGATCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 211
Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 212 AAAGAAGAACACAAAGACACACCGACACACAGTGGTTTCCTTTAGTGGGATCACTGAAC 271
Qy 81 ValAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
Db 272 GTAATAACGTTCTCGAGCCCGCAGCAAAACCAACCAATGAGAGAGAGAGAGAGAGAG 331
Qy 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
Db 332 CTGCACCAACAAATTTGAAATGATTAAGGAGCAAGTCAAGAGATGGGAGAGAAATCGCAG 391
Qy 121 GlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
Db 392 CAGCAGCAGACAGCAGAGGGTGTGCCCGACCTGTGGCATCTGCCACAGACAAATTT 451
Qy 141 AlaAspGlyCysGlyHisLysCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
Db 452 GCAGATGGATGGCGCCATTAATTTCTTATTCGCAAAACCAAGTTCTGTGCTCGATGTGA 511
Qy 161 GlyArgValSerLeuArgSerAsnLysValMetTyrValCysAsnLeuCysArgLysGln 180
Db 512 GGTGAGTGTCTTAGCTCAACCAAGTTATGTGGTGTGTAAATTTGGCCGAAACAA 571
Qy 181 GlnGluLeuLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
Db 572 CAAGAAATCCTCAATAATCAGAGCATGGTTTATATAATAGTGGGTCTAACACACTGCAG 631
Qy 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLysLys 220
Db 632 CAACCTGTATCAAAAGGTTCTCGAGGGCTTCGAAATGAGAGAGAGAGAGAGAGAGAG 691
Qy 221 AlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
Db 692 GCAAAACTACAGCAGCAGCCAGTTCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751
Qy 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrLysAsnGlySerGly 260
Db 752 GTTGAAAGGGCGAGCTCATGGGCTCACAAGCAGGATACTATTAAAAATGAGATCAGGA 811
Qy 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
Db 812 GTGAAGCACAGATTCAGATGACATGCCCTTCAGACAGAGAGAGAGAGAGAGAGAGAG 871
Qy 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
Db 872 AGGGATCAAAATCGAGATACAGCAAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAT 931
Qy 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320
Db 932 CCTTCAGATGGTACAAATGCCAAGATCTCTTCGGATTTATGCTGATAGACGATCTCAGCGT 991
Qy 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340

Db 992 GAGCCTCAATTTTATGAAGAACCTGGTCAATTTAAATTTACAGGGATTCTTAACAGGAGAGC 1051
Qy 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
Db 1052 CATAGACATTCCAAAGAGATATATTGTGATGATGAAGATGTGGAGACAGAGATGAATAT 1111
Qy 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
Db 1112 GAAGACAAAGGAGAGAGAGAGATACACGACCGCTACAGAGTGTATCAAAATCTCGCC 1171
Qy 381 ArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSer 400
Db 1172 CGGTATCCCGTAAAGCCACCAACCTACGAAGAAACAAATGGCATCCACGCTGAGGTGCC 1231
Qy 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAlaGluLeuGluAsp 420
Db 1232 AGGCGACGACATGAGAGAGAGCAAGTGTATTTCTTTGGCAACGCTGAACATAGAAGAT 1291
Qy 421 SerArgIleSerLeuLeuMetAspArgProSerArgGlnArgSerValSerGluArg 440
Db 1292 TCCAGGATTTCTCTGCTAAGGATGATAGACCATCAAGGCAAGATCTGTATCTGAACGT 1351
Qy 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
Db 1352 AGAGCTGCAATGGAAACCAACGATCGTATTCAATGGAAAGAACTCGAGAGGCTCAGGGA 1411
Qy 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
Db 1412 CAAGTTCCTTATCCACAAAGACCTCAATCATAGTCTCCACCCCTCGCGGAGGCCCT 1471
Qy 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
Db 1472 ATACCGTTGATAGACCAAGACATGAGGCGCGCTGACTCCCTACGGAAACAGACACCTTA 1531
Qy 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
Db 1532 GATCCAGCTCTGCTGTGAGGAAACCGAAGCGAGAAAAATGGAACCATGTGTAAGGAAT 1591
Qy 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProArgProHisLys 540
Db 1592 GATTTCTTTGAGTTTCAGACCTCGTCCGAGTCAGTGAGCGCCGCCCAACCAAGGCTCATAAA 1651
Qy 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAla 560
Db 1652 TCCAAAGAAAGAGGTAAATATGCCCGAGTTTCACTGACAGCTCGAGAGAGAGAGTGCA 1711
Qy 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
Db 1712 TCCACACCTGAGTATACAAAGCTGTGATGATGAGAGCTGGAAGCGAGAGTGTGAGTGAG 1771
Qy 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1772 AAAGGGGACAGTCAAAAGGGAAAAAGAAAACTAGTGAGCAGGGAGTTTGTGCGATCT 1831
Qy 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyLysHisSerLeuGluGlu 620
Db 1832 AACACAGGTCTGAGAGACAAAGAAAGAGATGATCTATGTTGGCCATCTCTTTGGAAGAG 1891
Qy 621 AspLeuGluTyrSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
Db 1892 GATTTGGAATGCTCTGAGCCTCAGATTAAGGACTCTGGGTAGATACCTGTAGTAGCACA 1951
Qy 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTyrGlnProSerLys 660
Db 1952 ACCCTTAACGAGGAGCATAGCCATAGTATAGCACCTCTGTGACCTGGCAGCATCCAAA 2011
Qy 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 2012 GATGAGATCGCCTCAATTTGCTGTATTTTATTAATAGCGTTTAAAGATGGGAGTGTA 2071
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyLysMetThrGluSer 700
Db 2072 CTTGAGACTCAGGAGCAATGCTGGCTTAAAGGTTTAGGAGAGAAAGATGACTGAATCA 2131

QY 701 GlyArgLeuCysAlaPheIleThrLysValLysGlySerLeuAlaAspThrValGly 720
DB 2132 GGTGGACTTTGTGCAATTATATACCAAGTAAAAAGAGATTAGCTGATCTGTAGGA 2191
QY 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
DB 2192 CATCTTAGACCAAGTGATGAAGTCTTGGAAATGGAGGCTATTTCGAAGGAGCCACA 2251
QY 741 PheGluGluValTyrAsnIleLeuLeuSerLysProGluProGlnValGluLeuVal 760
DB 2252 TTTGAGGAAGTTTACAACTATTCTAGAAATCCAAACCTGAACCAAGTTGAGCTTGT 2311
QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
DB 2312 GTTTCAGGCCAATTGGAGATATCTTAGAATACCTGATAGCAGCATGACAACTGGAA 2371
QY 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
DB 2372 TCCAGTTCTAGCTCATTTGAATCTCAAAAAATGGACCGTCTCTTATATCCGTTACCTCA 2431
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
DB 2432 CCCATGAGTCTCGCATGCTGAGGATGTCGCCAGATTCTTATCTGACAGCTTTCATA 2491
QY 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
DB 2492 AAATATGGTTTGAAGGTTGTGACCAAGTTGATAGTTTCAATTTTGGAGCAAGAGT 2551
QY 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
DB 2552 CTCCTCCAGGGAAGATGGAGGCCAAGAAATCCTTATGTAAATTTACTTCTTCCA 2611
QY 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
DB 2612 GATAGAGTGTATATAAATAAGAGAGAAACAAAAACAGTCAAGAAATCTTTGGAAACCCAAA 2671
QY 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
DB 2672 TGGAAACAGACTTTTCATTTATCTCTGTCACCGAAGAGAAATTCCTGTGAACGAATGCTG 2731
QY 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluLeuSerGluPheLeuGlyGlu 920
DB 2732 GAAATTTACCTTTGGGATCAAGCTAGAGTTCCAGAGAGAGAGCGGAATCTTAGAGAG 2791
QY 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisIleTyrLysLeuGln 940
DB 2792 ATTTTAATTGAATTTGAAACAGCTTTGCTAGATGATGAGCGCACCTGGTATAAGCTGCAG 2851
QY 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
DB 2852 ACCCATGATGTCCTCATTTGCCACTCCCTCGCCCTTCCCATATCTGCCCGGAGGAG 2911
QY 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
DB 2912 CTCATGGAGAGACCCCAACCGGAGGCTGCAAGAGTTCGAAGAAATAAGTCAGTGAA 2971
QY 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
DB 2972 GTGTCCTACGTACGACTCGAGATGGCGTGGAGTAGTGTCCAGATTATCGACACAATGGC 3031
QY 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
DB 3032 CGCATCTTCAAGCTCCACGTTGTGCGTGCACAACAAGTCATGTCAATAATCATTTGC 3091
QY 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
DB 3092 TCACCATCAGGGTCTCTCTCATCGAGTAGATGTTTAGGAAGACACAGGTCAATGTCGCT 3151
QY 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
DB 3152 AGTGCCCTCTCTCTCAAGGAATGTGAAACAGGGGCCACCGGGGACACGCTGCTACTGGC 3211

QY 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
DB 3212 CATTTACACACAAATTAGCCGAATGGATAGACACCGGTGTGATGATGCCACTACTCTTCA 3271
QY 1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
DB 3272 GATAGAGACAGGGAATTTGTGAAGCAGCAGATAGACAGCCATATCACAGATCCAGATCAACA 3331
QY 1101 GluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAsp 1120
DB 3332 GAAACAAGGCTCTCTCTAGAGCGGACCAACCCGCTCCAGATCCCTCTGAACGTCCTGAT 3391
QY 1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
DB 3392 ACAACCTCATGAGGTGATGCTTCAATTAAGTGAAGATCTGCCCTCTTCCCTTCCCT 3451
QY 1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
DB 3452 GCCTTATCAGAGTCTCACCTCGTACCGGTCTGTCCAGACAGCCCATCAAGTACTCCG 3511
QY 1161 GlyThrGlyArgArgGlyValArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
DB 3512 GGAACAGGACGAAGGGCGGACAGCTTCCACAGCTTCCACAAAGGGAACATTTGGAGAGA 3571
QY 1181 SerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
DB 3572 AGTGCTATGGATATAGAGAGAGAAATCGCCAAATGAACCTTAACAAATACAAACAGGTA 3631
QY 1201 AlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpAsp 1220
DB 3632 GCCGGATCAGACCCAGACTGGAGCAAGATTACCAITTCGAAGTATCGCTCAGGATGGAT 3691
QY 1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspVal 1240
DB 3692 CCACATAGAGGGCGAGATACTGTTTCCATTAATCTCCGACAGTGAATGAAGTGATGA 3751
QY 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerVal 1260
DB 3752 TCTCGGTTTCAAGGACTAGTAGTGTCTTCGTTTTCAGCAGCAGCACTACATGTCGCTC 3811
QY 1261 GlnSerGluArgProArgGlyAsnArgLysValIleSerValPheThrSerLysMetGlnAsn 1280
DB 3812 CAATCAGAGCGGCGGAGAGAAACAGGAAATCAGTGTCTTTACATCCAAAATGCAAAAC 3871
QY 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerLysSerGlyAspMet 1300
DB 3872 AGACAGATGGCGCTGTGGGGAAGAACTTCACCAAAAGCACCAAGCATCAGTGAGACA 3931
QY 1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320
DB 3932 TGCTCACTGGAGAGAAATGACCGGACGCTCCGACACTCGACTGGGCGCCCTGGGTACC 3991
QY 1321 SerGlyLysLysArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
DB 3992 AGTGGCAAGAACCGCGATCTAGCATTTGGGGCCAAAATGCTAGCTATTGTTGGTCTCTCA 4051
QY 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArg 1360
DB 4052 CGGAAAAGTCCAGTGCCTCTCAACTCAGCAAAACCGAAGGAGGAGTAAAAAGTACGG 4111
QY 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
DB 4112 AGCACTGTTTCAGAGAGACAGGAGACCGGCTAGAGTGGAGATGAGGAAGTGGATGACC 4171
QY 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn 1400
DB 4172 CGCCAGGCCAGCGGGAATCCACAGATGGCAGCATGAACAGCTATAGTTCGGAAGGAAT 4231
QY 1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
DB 4232 CTGATCTTCCCTGGGTCGCTCTGACAGCCAGTTCAGTCACTGATTTCTCTGGATGGC 4291
QY 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440

Db 3291 GATCATAGAACACAGAGATGCTAGTACTAGTATTATATCAGAACACAGACAGTCTCTT 3350
Qy 1088 AlalaAspArgGlnProTyrHisArgSerArgSerThrGlu 1101
Db 3351 ATGCTGCCAGACGA-----AAACAGGACGAAGTGCAGAA 3386

RESULT 3
US-10-240-965-72
; Sequence 72, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEILHAMER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAI, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 72
; LENGTH: 7325
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 977757.3
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 4103, 4411, 4430
; OTHER INFORMATION: a, t, c, g, or other
US-10-240-965-72

Alignment Scores:
Pred. No.: 7,94e-69 Length: 7325
Score: 1042.00 Matches: 245
Percent Similarity: 62.65% Conservative: 62
Best Local Similarity: 50.00% Mismatches: 113
Query Match: 12.58% Indels: 70
DB: 12 Gaps: 13

US-09-617-099B-1 (1-1590) x US-10-240-965-72 (1-7325)

Qy 1125 ArgSerMetPro-----SerLeuMetThrGlyArgSerAlaProSerProAla 1141
Db 66 CGGGCGATGCCCCGTGAGCTCCCTC-----GCCGCCCTCCCGCCCGCG 113

Qy 1142 ----LeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerThrPro 1160
Db 114 TGCCTATCACTCGGATCGCGCCAGCCTGGG-----GCCGGCGCGCCCTACTGCCG 167

Qy 1161 GlyThr-----GlyArgGlyArgGlnLeuProGlnLeu-----ProProLysGly 1176
Db 168 GGTTCGGGGGGGGGTCCTCGGGGACACCTCGCCCGCTTCGGAGCGCCCTCGG--- 224

Qy 1177 ThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLys 1196
Db 225 CCTGTGGAGGCCCTCCCTG-----245

Qy 1197 TyrLysGlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArg 1216
Db 246 -----TCTGGACCCCGGCCCCACCTCCGACCCCTTTTATCACAATCGC 287

Qy 1217 SerGlyTrpAspProHisArgGlyAlaAspThrValSerThrLysSer-----Ser 1233
Db 288 CTCCTCTGGAGCGCTGCCCTGATTGCCTTCACTCTTCTTGAATAATGGTGTTCG 347

Qy 1234 AspSerAspVal-SerAspValSer-----AlaValSerArgThrSerSe 1248
Db 348 AGAGTCAATTGAAGCCTTGCAAAATGCCCTAGGGGTGTCTGTGTGGAGGACGCCCCC 407
Qy 1248 rAlaSerArgPheSerSerThrSerTyrMetSerValGlnSerGluArgPro----- 1265
Db 408 TGTGATGCGGAACACACAGGCTCAGATTTCATGATTCGAGCTGCCTGAGCCCTCCACCC 467
Qy 1266 -----ArgGlyAsnArgLysIleSerValPhe-ThrSerLysMetGlnA 1280
Db 468 CGGGGACCATGTTTAACGGGAGCCAGGTCTCTGCTCATCTGGGGCTCC----- 517
Qy 1280 snArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerLysSerGlyAspM 1300
Db 518 -----AGGAATGTGTGCGGAGCTCCAGCATTTGCGGTGAA 554
Qy 1300 etCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyT 1320
Db 555 TCTGCGGATCCAGCAGCCGCGGGC-----GGGGCTGGGACCA 593
Qy 1320 hrSerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuS 1340
Db 594 CCACCGCCCAAGAGCGCGGAGCAGCTCGGTGTCAGAGATGCTGCCATCGTGGGCTGA 653
Qy 1340 erArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuA 1360
Db 654 CTCAGTGGAGCAAGAGCACACTCCAGCTTCCGACGCTTGAAGGGGCCCAAGAAAGCTGC 713
Qy 1360 tgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTpmMetT 1380
Db 714 GCAGCAACATCCGCGGAGCAGGAGACAGGATCGCGTGGAGATGCGAGCCCGGTCA 773
Qy 1380 hrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyA 1400
Db 774 CAGCCAGCGGAGCGGAGTCCACCGATGGGAGCACCAACACAGCAACAGCTCCGAGGCA 833
Qy 1400 snLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspG 1420
Db 834 CGTTTCATCTTCCCTACTACCCGCTAGGGGCTGAAAGCCAGTTCAGCGATTTCCTGGATG 893
Qy 1420 lyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspI 1440
Db 894 GGCTGGGACCATCAGATTGTGGGCGGACAGACACTGGCAACACACCCATCGGGAGATG 953
Qy 1440 leGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArgG 1460
Db 954 TGCACATTGCCATCATGGACCGGAGTGGCCAGCTGGAGTGGAGTGAAGTTCGCTCGGG 1013
Qy 1460 lyLeuValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuL 1480
Db 1014 GGCTGACCCCAACACAGGCTCCAAATCCTCCAGCCCACTATATCAAGTTTACCTGC 1073
Qy 1480 euAspAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeuGluP 1500
Db 1074 TGGAGATGGGCGCTGCTTGGCCAGAGAACAGAACAGAGTGAAGAGACTGTGATC 1133
Qy 1500 roLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleI 1520
Db 1134 CCTGTACAGCAGGCTCTGCTTTTGACAGGGGACCCCGGGCAGGTGCTGCGAGTGA 1193
Qy 1520 leValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleI 1540
Db 1194 TCGTCTGGGAGACATATGCGCATGAGCACACAGTGTCTTATGGGCATGGGCCAGATCA 1253
Qy 1540 euLeuAspGlnLeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPheProProS 1560
Db 1254 TGTGACAGCTGGACCTCAGCGCGCGGTCAACCGGTGGTACAAACTCTTCCCACTCT 1313
Qy 1560 erSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuG 1580
Db 1314 CCTCAGTGGAGACTCCACACTCGGATCCCTCACAGGGCGCTGTCCCACTCTTCCCTGG 1373

QY 414 -----AlaAsnAlaGluLeuGluAspSe 421
Db 2178 CCAGGTACAGTGGCGGTCTACGCTCCAGAACCCAGCCAGCGTGGCGGTCTACGCTC 2237
QY 421 rArgIleSerLeuLeuMetAspArg-----ProSerArg----- 433
Db 2238 TAGAACCCAGCTAGACGAGTGGTGGTCTACGCTCCAGAACACCCAGCCAGGAGGAGG 2297
QY 434 -----GlnArgSerValSerGluArgAlaAlaMetGluAsnG1 447
Db 2298 GTCTGGTCTAGGACACCAAGACGAGGAAGATCCCGCAGTAGAAGCTTAGTTAGAGTGG 2357
QY 447 nArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnAr 467
Db 2358 AAGATCTCACTCT-----AGAACCTCAAGAGAGGAGGAGTCTGGCTCATCTTCAGA 2411
QY 467 gThrSerAsnHisSerProProThrProArgArgSerProIleProLeuAspArgProAs 487
Db 2412 GCGGAAAAACAATCCAGAACATCTCAAGAGAGAGCAGG---TCCAATTCAGCCAGCA 2468
QY 487 pMetArgArgAlaAsp---SerLeuArgGlyGlnHisLeuAspProSerSerAlaVa 506
Db 2469 AATGAAGAAATCTCGCATTTCTTCAAGCGGAGCAGGTCTCTCTCTCCACCGGTCCAA 2528
QY 506 lArgIleThrLys-----ArgG1 512
Db 2529 AGCAAAATCTGGTCTTCTTTGAGCGCGACCTTTTCAGGGTCTTCCCCATGCCCTAAGCA 2588
QY 512 uLysMetGluThrMetLeuArgAsnAspSerLeuSerSerAspGln----- 527
Db 2589 AAGTCCAGACACCCAGCGGAGTGGTCTCGATCTCTCCACCTAAAGCTAAATC 2648
QY 528 -----SerGluSerValArgProProProArg----- 537
Db 2649 TAGAACGCCACCCAGAGCGAGTCTCTCCAGTCTTCTCGGCCACCTAAACAGAAATCTAA 2708
QY 538 -----ProHis-----LysSerIleLysGlyG1 545
Db 2709 GACACCATCAAGACAAAGTCATTCAGTTTCATCTCTCTCAATCTCTAAAGTGAATCTGGAAC 2768
QY 545 yLysMetArgGlnValSerLeuSerSerSerGlu-----GluGluLeuAlaSerThrProG1 564
Db 2769 ACCACCGAGCAAGGGTCCATAACAGTCCCGAGGCCAATGAGCAATCTGTAAACGCCACA 2828
QY 564 uTyrThrSerCys-----AspAspValGluLeuGluSerGluSer----- 577
Db 2829 GAGACGGAGCTGTTTGAATCATCATCCTGACCTGACCTGAGTTGAATCTAGGACCCCTTCTAG 2888
QY 577 ----- 577
Db 2889 ACATAGTGTCTCAGGGTCTCTCTCTCTAGAGTGAATCTAGCACACCTCCACAGACAGAG 2948
QY 578 -----ValSerGluLysGlyAs 583
Db 2949 CCATCTAGTGTATCATCTCCACAAACCAAGTGAAGGCAATATATCACAAGACAAAG 3008
QY 583 pSerGlnLysGlyArgLysThrSerGluGlnGlyValLeuSer----- 598
Db 3009 AAGCCATTCTGGCTCTCTCTTCCAACTCTCTAGTGGGTGACGTCGAGAACAACTCCACG 3068
QY 598 ----- 598
Db 3069 GCGAAGCAGATCAGTATCTCCCTGCTCCAATGTGGAATCCAGATTTGTGCCAAGATACAG 3128
QY 599 -----AspSerAsnThrArgSerGluArgGlnLysLysArgMe 611
Db 3129 TCATTCTGGGTCTCTCTCACCAGATACCAAGTGAACCTGAAACACCGCCAGACAAAG 3188
QY 611 tTyrTyrGlyGlyHisSerLeuGluGluAspLeuGluTrpSerGluPro----- 627
Db 3189 TCACCTCAGGTCTATTTCACCATACCCCAAGTAAAGGCCCAAACTCCACCGGGCCCAAG 3248
QY 628 -----GlnIleLysAspSerGlyValAspTh 636

Db 3249 TCTTTCTGGATCAAGTCAACATCACCATGTCCCAAGAGAGTCTAAAGACTCACTAGTTCARAG 3308
QY 636 rCys-----SerSerThrThrLeuAsnGluG1 645
Db 3309 TTGCCCTGGATCCCTCTCTCTCTGCGAGGAGTAAATCTAGCACACCCAGCGGCGAGAG 3368
QY 645 uHis----- 646
Db 3369 CTATTTTGGTGTCTCATCTCTGCAACTGAAAGGACAACTCAAACTCCAGACACCAG 3428
QY 647 -----SerHisSerAspLysHis----- 652
Db 3429 ATCTGATATCTCAAGTCCAGAGTCCAGACAGTCAATTCAGATCACCATCTCTCGCAGAG 3488
QY 653 -----ProValThr 656
Db 3489 CAAATCTCAACATCACCTAAGGAGGTGGTCCAGGTCTTCTCATCTCCAGTCACTGAGT 3548
QY 656 rGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuL 676
Db 3549 GGCATCCAGATCTCCAATAAGACAAGATAGA---GGTGAGTTCTCAGCGAGTCTCTATGTGA 3607
QY 676 yAspGlySerValProArg----- 682
Db 3608 AATCTGGATGTCTCTCTGAGCAGACAGGTTCCAGTCTGACTCTTCTCATATCTCTACAG 3667
QY 683 --AspSerGlyAlaMetLeuGlyLeuLysValValGly-----G 695
Db 3668 TGGACTCGAATTTCTCTTTGGGGCAGAGTAG---ATTGGAGACTGTGTAATCAAAAGAGAAA 3726
QY 695 lYysMetThrLysSerGlyArgLeuCysAlaPheIleThrLys---ValLysLysGlySer 714
Db 3727 ATGGCTTTACCCCTCAGGAGGA---TGCTACTGCATCACCCTCTCTAGACAGAAAGACAAA 3783
QY 715 LeuAlaAspThrValGlyHisLeuArgPro----- 724
Db 3784 TTTAGTCCCTTCCAGTACAGGATAGGCTGAGTCTTCACTGGTATTCAAAGACACACTT 3843
QY 725 -----GlyAspGluValLeuGluTrpAsn 732
Db 3844 AGAACCCCGCCCAAGAGAAAGTGTGTGGTGCATCTCCAGAACAAAGAGCAAAAT 3903
QY 733 GlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGluSerLys 752
Db 3904 AGTGCAATTCCTACGTCAAGCAGATGAAGAGTTAATGGAGTGGTAGAGTCTGAA 3963
QY 753 ProGluProGlnValGluLeuValSerArgProIleGlyAspIleProArgIlePro 772
Db 3964 -----GNAACCGCAGGCCAAATC-----CTGTCT 3987
QY 773 AspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSer-----GlnLys 790
Db 3988 CATTTGTCTCAGAACTTAAAGAAATGTCCACAGTAATCTTGAATCATCTCTCTGAAGTA 4047
QY 791 MetAspArgProSerLysSerValThrSerProMetSerProGlyMetLeuArg----- 808
Db 4048 GAAGAAAGCCCTGCTGTCTTGTACTCTTGATCAGACAGCAGTCACAGGCTCTTTTGGAA 4107
QY 809 -----AspValProGlnPheLeuSerGlyGlnLeuSerIleLysLeuTrp----- 823
Db 4108 GCGTAGAGTCTCTTCAATGGCTCATCT-----TGGGTGGGCCA 4149
QY 824 ---PheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuPro 842
Db 4150 CATTTTCTCCAGAACATAAAGAACTG-----TCTAACTCCCA 4188
QY 843 SerArgGluAsp-----GlyArgPro-----ArgAsnProTyrValLysIleTyr 857
Db 4189 CTGAGGAGAACAGCTTTGGATCACCTTAGAATTTAGAACTTCAGGC----- 4236
QY 858 PheLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeu 877

Db 4237 -----CCACTTGGTACAGAAATGAATACCTGATTTCTCTGAGTTAAAGAAATTTG 4290
Qy 878 GluProlystPAsnGlnThrPheIleTyrSerProValHisArgGluPheArg--- 896
Db 4291 -----AATGGACCTTTCTTAA--TCAGCTGGAACAGATCCATCTTAGACAT 4337
Qy 897 GluArgMetLeuGluIleThrLeuTrpAspGln----- 907
Db 4338 GAAAGAACAAATCGACAGATCTCTGGACACAGCAGTCTGAGTTATCCCCAGATGCAGT 4397
Qy 908 AlaArgValArgGlu--GluGluSerGluPheLeuGlyGluIleLeuIleGluLeuGluTh 927
Db 4398 GGAAGAGCGAGGATGCTTCAAAATCAG-----AGCATCTCTTC 4436
Qy 927 rAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGlnThrHisAspValSerLe 947
Db 4437 ACCTGTGCTGATGCTGTACCCAGAACCCCTCGAGAGAAAGATAGTCTTCGATCTTC 4496
Qy 947 uPro-----LeuProArg---ProSerProTyrLeuProArgGlnLe 961
Db 4497 TCCTGAATGAAGATGTTTACCAGAACTCCATCA-----AGGAGAGCAG 4544
Qy 961 uHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGluVa 981
Db 4545 GTCTGGGTCTCTCCAGACTTAGA----- 4569
Qy 981 lSerAspTyrAspCysGluAspGlyValValSerAspTyrArgHisAsnGlyAr 1001
Db 4570 -----GATGGTCTGGGACTCCCTCG-----AGGCACGCTCTC 4604
Qy 1001 gAspLeuGlnSerSerThrLeuSerValProGlu--GlnValMetSerSerAsnHisCy 1020
Db 4605 TGGGTCTCTCTGGAATGAAGATATACCTAGACGCCATTAGAGGAGAGCGAATG 4664
Qy 1020 sSerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPr 1040
Db 4665 TGATTTCTCCAGAACCGAAGCTTTGCCTCAGACTCTTAGCCGAGAGTGTCTCTCC 4724
Qy 1040 oSerAlaPro-----ProGlnArgAsnValGluGlnGlyHi 1053
Db 4725 ATCATCCCGAGAGCTCAACAAAGTGTCTTACCCCGCAGAGAA-----AG 4772
Qy 1053 sArgGlyThrArgAlaThrGlyHisTyrAsnThrIle----- 1065
Db 4773 AAGCGGGTCAGATCATAGTTGATCAGAAACTGTGCTCGACTCCCTCGGGGCAGAG 4832
Qy 1066 -SerArgMetAspArgHisArgValMetAsp-----AspHi 1077
Db 4833 AAGTCGTTCCGGATCTCTCAAGAACTTGATGTGAACCCAGTGCATCCCTCAGGAAG 4892
Qy 1077 sTyrSerSerAspArgAspArgCysGluAlaAlaAspArgGlnProTyr---HisAr 1096
Db 4893 AAGTGAGTCAGACTCTCTCCAGATTTCTAAAGCCCAAGACAGACCCCACTTCGGCAGAG 4952
Qy 1096 gSerArgSerThrGluGlnArgProLeuLeuGluArg-----ThrThrThrArgSe 1113
Db 4953 GAGTCGGTCTGATCATCTCCAGAGTTGACGCAAACTCGACTATCCCTCCGGCGAG 5012
Qy 1113 rArgSerSerGluArgProAspThr---AsnLeuMetArgSerMetProSerLeuMetTh 1132
Db 5013 TAGGTCTGTTCTCTCCCTGAGTGAAGATAAGCCAGACAGCAGCACCAGCGGCACAGAG 5072
Qy 1132 rGlyArgSerAlaProProSerPro-----AlaLeuSerArgSerHisProArg----- 1148
Db 5073 TGGTTCTGATTCCTCTCTGAACTTAAAGCTCCAGCCCTCGGGCCCTTCCAGACGAGAG 5132
Qy 1149 -----ThrGlySerValG1 1153
Db 5133 CAGATCAGGTTTCATCAAGCAAGGAGGAGCCCTTCTCTGAAAGAGCAGCAGTACCAG 5192
Qy 1153 nThrSerProSerSerThrPro-----GlyThrGlyArgArgGly---ArgGlnLeuPr 1170
Db 5193 GTCTCTCTCTGAACTCCGCCCCCAAAATCCAGAACTCTCTCGAGAGGTTCCAGTCTATCACC 5252

Qy 1170 oGln-----LeuProProLysGlyThrLeuGluArgSerAlaMetAs 1184
Db 5253 AGAGCCCAAGACCAAGTCTCGTACACCACTCGA-----CGTCCAGCTCTCCATC 5303
Qy 1184 pIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGlySerAs 1204
Db 5304 ATCTCCGGAGCTAAACAAGGAAGCCAGCTGTCCGTAGAAGCCGCTCTGCTCATCTCCT 5363
Qy 1204 pProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpAspPro---HisAr 1223
Db 5364 ACCA-----GAAACTCGCTCTAGAACTCCCCCAAGGACCG 5399
Qy 1223 gGlyAlaAspThrValSerThrLysSerSerAspValSerAspValSerAlaVa 1243
Db 5400 GAGAGTCTCTCAGTGTCTTCC-----CCGAGCCAGCCGGAATAATCGAGGTCTTCACCGC 5456
Qy 1243 lSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSer---ValGlnSe 1262
Db 5457 ACGCGCTCAGTTCATCTCCAGCACTAAGACAACCTCAAGGAGAGCGCGCTCTCTCTTC 5516
Qy 1262 rGluArgProArgGly-----AsnArgLysLysSerValPheTh 1275
Db 5517 GCCAAAGCTCTGTGACTCCAGAGGTCCCGTTCCTCCGCTCAAGGAGAGAGAAAACAAGA 5576
Qy 1275 rSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSe 1295
Db 5577 AACCCGACGTCGAGATAGTCTGATCTTCTAGTCAACCTCTCGGGAAGACAGCGGAG 5636
Qy 1295 rIleSerGlyAspMetCysSerLeuGluLysAsnAspGly-----SerGlnSe 1311
Db 5637 CCGGTCAAGTTCGCGGTACTCGCGCGGAGGGAGGCTCTGTTATCATCTCAAGTTC 5696
Qy 1311 rAspThrAlaValGlyAlaLeuGlyThrSerGlyLysArgArgSerIleGlyAl 1331
Db 5697 ACCTGCCCGGAGCAAAAGTTCCCGACCTCTCTCGACCGCGAAGA----- 5742
Qy 1331 aLysMetValAlaIleValGlyLeuSerArg-----LysSerArgSerAlaSerG1 1348
Db 5743 -----GGCCCTCTCGAGACACCCCAACCCAGTCCGAAGCGTTCTTCG 5783
Qy 1348 nLeuSerGlnThrGluGlyGlyLysLysLeuArgSer-----ThrVa 1363
Db 5784 CTCAGGCACATCACCCAGCCCGTGGAAACGCTCTAGATCTCGAGCTCTCCAGCCACTCA 5843
Qy 1363 lGlnArgSer---ThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgG1 1382
Db 5844 CCGCGCATCCAGGTCCAGAACCCCTCG-----ATAAGCCGACG 5882
Qy 1382 nAlaSerArgGluSerThr 1388
Db 5883 TAGGTCCAGATCTCGAACT 5901

RESULT 5

US-09-764-891-9844/c
; Sequence 9844, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9844
; LENGTH: 24387
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-9844

Alignment Scores:

Pred. No.:	3,36e-17	Length:	24387
Score:	368.50	Matches:	354
Percent Similarity:	33.15%	Conservative:	180
Best Local Similarity:	21.97%	Mismatches:	548
Query Match:	4.45%	Indels:	532
DB:	11	Gaps:	78

US-09-617-099b-1 (1-1590) x US-09-764-891-9844 (1-24387)

QY	193	AsnSerGlySerAsnThrLeu-GlnGlnProAspGlnLysValProArgGlyLeuArgAs	212
DB	15016	AACTCGGACCGTTCCACCACTAAGTCTCC--GAGAACTTCCCGAGTCTTCTTCCTC	14960
QY	212	nGluGlu-----AlaProGlnGluLysLysLysLeuHis	224
DB	14959	AGAGAGCAGCCACCATCCCTCACTACCAAGTTCTTCGGCATGCCAGCTTTCCTCC	14900
QY	225	-GlnGlnProGlnPheGlnGlyAlaProGly-----AspLeuSerValProAlaVa	241
DB	14899	AGAAAGTCTTAACCTGCTCCAGCTCCAGGGTCCACCGAGAGATTCTTCTTCCCCAC	14840
QY	241	LGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrLysAsnGlySerGlyVa	261
DB	14839	ATCTAAGATCGTCACTAGCGCGAGCAAAACGGATAAATCATCTTCATACCCCTTC	14780
QY	261	lLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSerAr	281
DB	14779	CGGTAGGATGGGAGGTCCCGTAGCCCTGCCACCGCTAAGAGAGGGCGATCTGGGTCTG	14720
QY	281	sAspGlnAsnArgAtqTyrrGluGlnSerGluArgGluAspTyrrSerGlnTyrrValPr	301
DB	14719	AACCCCTACACAGAGA---GGTCATCTCGATCCCGATCTCCCGAGTGGGTAGTCCAG	14663
QY	301	oSerAspGlyThrMetProArgSerProSerAspTyrrAlaAspArgArgSerGlnArgL	321
DB	14662	GTCTGCACAGAGTGGGGAAGATCTAGAAGCCCCCAGCGAGTGGCGCTCT---AGGTC	14606
QY	321	uProGlnPheTyrrGluGluuProGlyHisLeuAsnTyrrArgAsp-----	335
DB	14605	TCCTCAG-----CGACAGGCTGGTCTTAGGACAGAAATATACCAGAGAGAGCGAG	14555
QY	336	-----SerAsnArgArgGlyHisArgHisSerLysGluTyrrIleValAspAspGluAs	353
DB	14554	GTCTAGTCAAGCGAGGAGGAGTCCCATCTAGATCCCGCAGCAGTGGGTAGATC	14495
QY	353	pValGluSerArgAspGluTyrrGluArgGlnArgArgGluGluGluTyrrGlnAlaArgTy	373
DB	14494	T---CGTCTTAGAACACACCGCCCGGGGCGAGTCCCGCTCTAGAACACCTCCAGGCG	14438
QY	373	rArgSerAspProAsnLeuAlaArgTyrrProValLysProGln-----	387
DB	14437	GAGATCA-----CGATCCAGAACTCCACCGCGGTAGGTCTCGGTCTAGAACACC	14387
QY	388	-----ProTyrrGluGlnMetArgIleHisAl	397
DB	14386	AGCCCGAGGGGAGGTCTCGGTCTAGAACACCTCTAGCGGCGAGATCTAGGACCCGATC	14327
QY	397	aGluVal-----SerArgAlaArgHisGluArgArgHisSerAspValSerLeu--	413
DB	14326	ACCAGTACGACGCGAGGTCTCGTAGTAGATCACCAGCCAGGAGAGTGGCAGGTCA	14267
QY	414	-----AlaAsnAlaGluLeuG1	419
DB	14266	TAGAACCCAGCTAGAGTGGCGCTCACGCTCCAGAACCCCGCAGCGTGGCGCGCTC	14207
QY	419	uAspSerArgIleSerLeuLeuArgMetAspArg-----ProSerArg--	433
DB	14206	ACGCTCTAGAACCCAGCTAGACGAGTGGTCTCGCTCCAGTCCAGAACACACGACGAG	14147
QY	434	-----GlnArgSerValSerGluArgArgAlaAlaMetG1	445
DB	14146	AGGAGGTCTCGGTCTAGGACACCAAGACGAGGAGATCCGCGAGTAGAAGCTTAGT	14087

QY	445	uAsnGlnArgSerTyrrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSerTyrrPr	465
DB	14086	ACGTGGAGATCTCACTCT-----AGAACACCTCAAGAGAGCGAGATCTGGCTCATC	14033
QY	465	oGlnArgThrSerAsnHisSerProProThrProArgArgSerProIleProLeuAspAr	485
DB	14032	TTCAAGCGGAAACAAATCCAGAACATCTCAAGAGAGAGCAGG---TCAATTCAG	13976
QY	485	sProAspMetArgArgAlaAsp---SerLeuArgLysGlnHisHisLeuAspProSerSe	504
DB	13975	CCGAGAAATGAGAAATCTCGCATTTCTTCAAGCGGAGCAGGTCTCTCTTCCACCAG	13916
QY	504	rAlaValArgLysThrLys-----	510
DB	13915	GTCCAAAGCAAAATCTCGTTGTCTTTGAGGCGCAGCTTTTCAGGTTCTTCCCATGCC	13856
QY	511	-ArgLysMetGluThrMetLeuArgAsnAspSerLeuSerSerSerAspGln-----	527
DB	13855	TAAGCAAAAGTCACAGACACACCCAGCGCGAGTCTGATCTCTCCAACTTAAAGC	13796
QY	528	-----SerGluSerValArgProProProProArg-----	537
DB	13795	TAAATCTAGACGCCACCCAGACGCGAGTCTTCTTCGCGCACCTTAAACAGAA	13736
QY	538	-----ProHis-----LysSerLysLy	543
DB	13735	ATCTAAGACACCATCAAGCAAAAGTCATCTCCAGTTCATCTCTCTCTTAAAGTGAATC	13676
QY	543	oGlyLysMetArgGlnValSerLeuSerSerSerGlu---GluGluLeuAlaSerTh	562
DB	13675	TGGAACACACCGAGCGAAGGTCTCAATAACAGTCCCGAGCCCAATGAGCAATCTGTAA	13616
QY	562	rProGluTyrrThrSerCys-----AspAspValGluLeuGluSerGluSer---	577
DB	13615	GCCACAGACAGCGAGTGTTTGAATCATCATCTCACCTCGCTGAGTTGAAATCTAGAC	13556
QY	577	-----	577
DB	13555	TTCTAGACATAGTCTCAGGGTCTCTCTCTCTCTAGATGTAATCTAGCACACCTCCAG	13496
QY	578	-----ValSerGluLy	581
DB	13495	ACAGAGCCCATCTAGTCTCATCTCCACACCCAAAGTGAAGGCAATAATATATACCAAG	13436
QY	581	oGlyAspSerGlnLysGlyArgLysThrSerGluGlnGlyValLeuSer---	598
DB	13435	ACAAAGAGCCCATCTGGCTCTCTCTTCTCAAGTCTTAGTAGGGTGAOCGTGAGAACAC	13376
QY	598	-----	598
DB	13375	TCCACGGCGAAGCAGATCAGTATCTCCCTGCTCCAATGTGGAATCCAGATTGTGCCAAG	13316
QY	599	-----AspSerAsnThrArgSerGluArgGlnLysLy	609
DB	13315	ATACAGTCAATCTGGTCTCTCTCCACAGATACCAAGTGAACCTTGAACACCCGCAAG	13256
QY	609	sArgMetTyrrGlyGlyHisSerLeuGluGluAspLeuGluTyrrSerGluPro----	627
DB	13255	ACAAAGTCACTAGGGTCTATTTCACCATACCCCAAGTAAAGGCGCCAACTCCACCGG	13196
QY	628	-----GlnIleLysAspSerGlyVa	634
DB	13195	GCCAACTCTTCTGGATCAAGTCACCATGTCCCGAAGAGAGTCTAAGATCTCATAGT	13136
QY	634	lAspThrCys-----SerSerThrThrLeuAs	643
DB	13135	TCAAAGTTCCCTGGATCCCTCTCTCTGTGCGAGGAGTAAATCTTAGCACACACAGG	13076
QY	643	nGluGluHis-----	646
DB	13075	CGAGAGCTATTTTGGTGTCTCATCTCTGCAACTGAAAGGCAATCTCAAACCTTACCAG	13016
QY	647	-----SerHisSerAspLys-His----	652


```
Qy 1222 -HisArgGlyAlaAspThrValSerThrLysSerSerAspSerValSerAspValSe 1241
Db 11044 GCACCGAGAGAGTCCCTCAGTGTCTTCC--CCGAGCCAGCCCGAATAATCGAGTCTTC 10988
Qy 1241 rAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerThrMetSer--Va 1260
Db 10987 ACGCGGAGCGGCTCAGCTTTCATCTCCACGCACTAAGACAACTCAAGGAGGCGCGTC 10928
Qy 1260 lGlnSerGluArgProArgGly-----AsnArgLysIleSerVa 1273
Db 10927 TCCITCGCAAGCTCTGTGACTCCAGAGGTCCCGTTCGCTCAAGGAGAGAGAAAC 10868
Qy 1273 lPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSe 1293
Db 10867 AAGAACAACCCGACGTCGAGATAGTCTTGATCTTCTCAGTCAACCTCTCGCGCAAGACA 10808
Qy 1293 rThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGly-----Se 1309
Db 10807 GCGGAGCCGCTCAGTCCGGGTACTCGCGCGCGGAGGAGGCTCTGGTTATCACTC 10748
Qy 1309 rGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSerSerIl 1329
Db 10747 AAGGTCACTCGCGGAGGAAAGTTCCCGGACCTCTCTCGACGCGAAGA----- 10696
Qy 1329 eGlyAlaLysMetValAlaIleValGlyLeuSerArg-----LysSerArgSerAl 1346
Db 10695 -----GGCGGCTCTCGGACACCCCAACCCACGTCGGAAGCG 10661
Qy 1346 aSerGlnLeuSerGlnThrGluGlyGlyGlyLysLysLeuArgSer----- 1361
Db 10660 TTCGCTGCAGCAGCATCACAGCCCGCTGGAAGCTCTAGATTCGAGCTCTCCAGC 10601
Qy 1362 -ThrValGlnArgSer---ThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetTh 1380
Db 10600 CACTCACCGGAGATCCAGGTCAGAAACCCCTG-----ATAAG 10562
Qy 1380 rArgGlnAlaSerArgGluSerThr 1388
Db 10561 CCGACGTAGTCCAGATCTCGAACT 10537
```

RESULT 6

```
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1
Alignment Scores:
```

```
Pred. No.: 8,12e-10 Length: 9025608
Score: 312.00 Matches: 420
Percent Similarity: 32.41% Conservative: 209
Best Local Similarity: 21.64% Mismatches: 696
Query Match: 3.77% Indels: 622
DB: 14 Gaps: 87

US-09-617-099B-1 (1-1590) x US-10-156-761-1 (1-9025608)

Qy 4 ProLeuGlyProArgGlyArgProAlaProThrPro-AlaAlaSerGlnProProG1 23
Db 8206041 CCCCAGGAGGAGGCGCGCGCGCGATCCACCTCCGAGGAGGCGCTCTCTCCCGCGCGAG 8205982
Qy 23 nProGluMetProAspLeuSerHisThrGluGluGluArgLysIleLeuAlaVa 43
Db 8205981 ATCCGCGCGCGCGAGCATGGCCACCC-GTCCGAGGAGCAGGCGCGCGCTACCCGCCCT 8205923
Qy 43 lMetAspArg---GlnLysLysGluGluLysGlnLysValLeuLysIleLysG1 62
Db 8205922 CGGCGACCGGCTCGGCGCGCGGTACTCACACGAGGCGGCGCGCGCGAGCGGTGT 8205863
Qy 62 uGluHisLysAlaGlnProThrGlnTrpPheProPheSerGly----- 76
Db 8205862 CGAGCACTCTGTCAAGCCC-----GGAGGTGCGCACCGCGACGCGCG 8205821
Qy 77 -----lIleThrGluLeuValAsnValLeuGlnProGlnLysGlnPro-AsnG 94
Db 8205820 GGAACGACACCGACCGAGCTCTCTCGGCGCTCCGCGCGCGCGCGCGCGCGCGCG 8205761
Qy 94 lLysGluPro-----GlnThrLysLeuHisGlnGlnPheG 106
Db 8205760 GGAACCCACCGACTCCGGACAGTCCAGCGCGGTCTCTCCGTCGCGCACCAACCGCAC 8205701
Qy 106 lMetTyrlLysGluGlnValLysLysMetGlyGluGluSerGlnGln----- 122
Db 8205700 GCAGCGCGCGCGCTCTGTGCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 8205641
Qy 123 -----GlnGluGlnLysGlyAspAlaProThrCysGlyLleCysHisLysT 138
Db 8205640 CCGCGCACTGCGCGCAGCGCATCCGGGGCAGTAAACCCCGCGCGCGCGCGCGCGCGCG 8205581
Qy 138 hrLysPheAlaAspGlyCysGlyHisAsnCysSerTyrlCysGlnThrLysPheCysAlaA 158
Db 8205580 CG-----GTCCTGGCGCGACCGCTCCCGCAGCGCGCGCGCGCGCGCGCG 8205539
Qy 158 rgCysGlyGlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysA 178
Db 8205538 GGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8205507
Qy 178 rgLysGlnGlnGluIleLeuThrLysSerGlyAla-----TrpP 191
Db 8205506 --TCCCG-CCGAGATCCCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8205450
Qy 191 heTyrlAsnSerGlySerAsnThrLeuGlnGlnProAspGlnLysValProArgGlyLeu- 210
Db 8205449 TGTTCGCGCTCGGCGCAGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8205390
Qy 211 --ArgAsnGluGluAlaProGlnGluLysLysAlaLysLeuHisGlnGlnProGlnPheG 230
Db 8205389 ACCGTGCAGTCCAGCTCCCGCAGCAAAACCCGCACTTGTCCCGGTGCGCGCGCGCG 8205330
Qy 230 lNglyAlaPro-----GlyAspLeuSerVal- 238
Db 8205329 AGCACCTCCGCACATGGGAAACGCGCGGTGTGTGCTGTGTGGAGCATCCCGCGTCG 8205270
Qy 239 -----ProAlaValGluLysGlyArgAlaHisGlyLeuThrArgGlnA 253
Db 8205269 TCCACAGCACAACAGCGCGCGAGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8205210
Qy 253 spThrIleLysAsnGlySerGlyValLysHisGlnIle-----AlaSerAspMetProS 271
Db 8205209 ACCACCGCTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8205151
```


Db 8203165 GCGTGGCAGACACAGCGTGGCGAGGTGGATCGCAGCGCATCAGCGTGGCGTCTCGTC 8203106
Qy 842 roSerArgGlu-----AspGlyArgProArgAsnProTyrValIleTyrP 858
Db 8203105 CGTCGGTAGCGGAACATCAGGCGCTGGTGGCGCGCCCTCGCGT----- 8203059
Qy 858 heLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValIysLysThrLeuG 878
Db 8203058 -----CCAGCTCGTCTCGTCCCTCGACCGCTCTCTAGCGCGTGTGCACACCCCTG 8203004
Qy 878 luProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArg-Glu 897
Db 8203003 CCGCATGTCGGGGACTCGAGCCGATCGACACCGACCGCCGACGAGAGCTCCGTCGAA 8202944
Qy 898 ArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPhe 917
Db 8202943 GCGCTTCTTCGAGAGTCGTAGCGCAT-----CTCCAGGATCTTGTGCGGACCCAG 8202893
Qy 918 LeuGlyGluIleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyr 937
Db 8202892 CTGGCGCATCGGTGGTACGCCTTGTGTGTGACCTCGCCGCG----- 8202851
Qy 938 LysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuPro 957
Db 8202850 -----CAGTGCACAGCCCGTCT----- 8202830
Qy 958 ArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSer 977
Db 8202829 GATCAACGTCCTCACCGCGAC---CCGGATGTGCGGTCTCTACGCGAAGCGCATCAAG 8202773
Qy 978 AspSerGlu-----ValSerAspTyrAspCysGluAspGlyValGlyVal-Valse 994
Db 8202772 AATGTTGTCTGATCTGTGTACGCGATCTT-----GTCCGGGTGGCCCTCGGTCTAC 8202722
Qy 994 rAspTyrArgHisAsnGlyArg----- 1001
Db 8202721 GACTCCGAGGTGAACAGCGACGCGACACACGCTCCCTGGGTGGACGGCTGCTCG 8202662
Qy 1002 ----AspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCy 1020
Db 8202661 CTGATCATTTGGCGACGGAGCGAGGTGCGCCCGCGTCTGTC----- 8202619
Qy 1020 sSerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPr 1040
Db 8202618 -----CGAGAACAGTTTATCGGTGCGACTCG-----CC 8202590
Qy 1040 oSerAlaPro----- 1044
Db 8202589 ATCCGCCCCCTGTCTCGCTCTCGGGAGGCTGTGACCTGCGGACGGGCACTTCGCCC 8202530
Qy 1044 oProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnTh 1064
Db 8202529 AATGCCACGGGTCTCCGCGAGCTCACCACACCCGCTTTCGAGCGATTTCGAGACGCC 8202470
Qy 1064 rIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSerAspArgAspAr 1084
Db 8202469 TTTGGCGAATGGGACCGTCTAC-----CCGAGCCGGGGCAGCAGCATCCCA 8202422
Qy 1084 gAspCysGluAla---AlaAspArgGlnProTyrHisArgSerArgSerThrGluGlnAr 1103
Db 8202421 CACGTTCTCGGCCCAAGCGCTCTTGGCCCGGTCGAGCAGCGGATCTCGTTCGTCGCG 8202362
Qy 1103 gProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAspThrAsnLe 1123
Db 8202361 GCCG-----AGCACACGCGCTCTGTTCTCTCCGAG-----CCGAAAGT 8202323
Qy 1123 uMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAlaLeuSe 1143
Db 8202322 CTTGCGCTCCCGCACCTCGTTCCACACGAGAGGTGCGAGCCCTTGGCTGCCAGTTCGT 8202263
Qy 1143 rArgSerHisProArgThrGlySerValGlnThrSer---ProSerSerThr---ProGl 1161
Db 8202262 ACGGCGCTTGGCAGGACGTCTGCTCGTCTCGCGCGGAGAGCCGACGACCATCTGACCGGG 8202203

Qy 1161 yThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArgSe 1181
Db 8202202 GCGCGCGCGTCTC-GGCCGAGATCTC-----CGGA 8202174
Qy 1181 rAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLys-----TyrLysGl 1199
Db 8202173 GAATGTCGGATTACGACACAGACGATCGTTCGGCTCTCTGCCGCTCTCTCTCTTGA 8202114
Qy 1199 nValaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTr 1219
Db 8202113 TCTTTCGGTCTCGGTACGTGCG-----CGACGGAAGTCCGCCACGCGCG 8202069
Qy 1219 pAspProHisArgGlyAla----- 1225
Db 8202068 CGCCATCACCGCGCTCGGCTCCGAGGCGCTTCAGCACCGCTTCGCGAGTTGCA 8202009
Qy 1226 -----AspThrValSerThrLysSer-SerAspSerAspValSerAspV 1240
Db 8202008 CCGCGCTCCGACCTGGAGACGTCACACCGCAGGTGGGAGCGCGGTTCGCGG 8201949
Qy 1240 alSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerV 1260
Db 8201948 CGACCAAGGTACCCGCGCGCGCGCGGTGCGGCGGTGCGGAGGCGGTACCCCTGCT 8201889
Qy 1260 alGlnSerGluArg---ProArg-----GlyAsnArgLysIleSerV 1273
Db 8201888 TGCCGAGAGCGGTGTCGAGAGACGCGGTCGAGGCTCGCGCGTACCGCGG 8201829
Qy 1273 alPheThr-----SerL 1277
Db 8201828 CGCTGACACACACGTCGCGCGCGGAGTTCGCGCGCGCGCGCGGACACCC 8201769
Qy 1277 ysMetGlnAsnArgGlnMetGlyVal---SerGlyLysAsn---LeuThrLysSerThrS 1295
Db 8201768 GCGCGCAGACCTCGAAGATCTCCGTGGGTGCGGCGAGCGGCGCTTTCGGGTGTCAGCG 8201709
Qy 1295 erIleSer-----GlyAspMetCysSerLeuGluLysAsnAspGly----- 1308
Db 8201708 CGGTGACGCGCGCCACCGCGGCTCGATGACGAGCGCGCGCGCGCGCGGACGCG 8201649
Qy 1309 -----SerGlnSerAspThrAlaValGlyAlaLeuGlyThrS 1321
Db 8201648 CQTTCCTCTGGGTGCGCGGTGTTCCACATCTCGTGTGATGCGGTGCGAGACGA 8201589
Qy 1321 erGlyLysLysArgArgSerSerIleGly----- 1330
Db 8201588 CCGACAGCGTGGGTGACGAGCGTTCGTGAGCAGGTCTGTCGCGGAGCGCGTGGCGG 8201529
Qy 1331 -----AlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnL 1349
Db 8201528 CCTTCGCGAGCATGTGCGCGGTGGCGGG---GCCACGACGACGAGTTCGGGTGCTGCGC 8201472
Qy 1349 euSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerThrGluT 1369
Db 8201471 CGATCGGACGTGCGGCACTCG-----TGACGCTGTCCTCCAGACCTCGGTGAGA 8201421
Qy 1369 hrGlyLeu----- 1371
Db 8201420 CCGGTTGCCGAGAGCGCGGACGAGTGGCGCGCGCGAGTGCAGCGCGCGGCGG 8201361
Qy 1372 -----AlaValGluMetArgAsnTrp----- 1378
Db 8201360 TGGGCACGACGCGAGTGTGTCCTCGTCCGATCTCGTCCAGCTTCGCGACGAGCTCACAGGCT 8201301
Qy 1379 -----MetThrArgGlnAlaSerArgLysSerThrArgLysSerM 1392
Db 8201300 TGAGCGCGCATGCGCCACTGATCCCGACGACGACCTTCGGTTCGCGGTCTC 8201241
Qy 1392 etAsnSerTyrSerSerGluGlyAsnLeuIle----- 1402
Db 8201240 CCCACACTCGTGAATCTACGGCTCGATGCTCAACATCGCGAGGGGAGCCACCGCGC 8201181

```
QY 1403 -----PheProGlyValArgLeuAlaSerAspSerG 1413
Db 8201180 CCGCGCACACACAGCCGCGAGTCGCGCGCTGAGATGAGCGAACAATGAGC 8201121
QY 1413 InPheSerAspPhe-----LeuAspGlyLeuGlyProAlaGlnLeuValGlyA 1429
Db 8201120 GAACGAACGC-CTACTGCGCGCGCGCTCGACGGGCTCGGA-----GGTC 8201077
QY 1429 rGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysG 1449
Db 8201076 AGCAGACCGCGTGTATCTCGCGAGCGCGAT----- 8201045
QY 1449 lYglnLeuGluValGluIleAlaArgAlaArgGlyLeuValValLysProGlySerLysT 1469
Db 8201044 -----CGAGAGCGG----- 8201036
QY 1469 hrLeuProAlaProTyrValLysValTyrLeuLeuAspIleValCysIleAlaLysL 1489
Db 8201035 -----CTTCTCGTGGAGCTGGGTGTC-----GACG 8201011
QY 1489 ysLysThrLysValAlaArgLysThrLeuGluProLeuTyrGlnLeuLeuSerPheG 1509
Db 8201010 AGGGGACCGAGCTCTGAGGAGCGCTTACCAGCTG----- 8200973
QY 1509 luGluSerProGlnGlyArgValLeuGlnIleIleValTTPGlyAspTyrGlyArgMetA 1529
Db 8200972 -----CGAGTAGTACGGCTGTGATCTGACGGCGCCCTGCGCGGTAG 8200930
QY 1529 spHisLysSerPheMetGlyValAlaGlnIleLeuLeuAspGluLeu 1544
Db 8200929 ATCAGAGGCTGTACTTCGAGTCAGTGGC-----CTCGAGGAGCTC 8200889
```

RESULT 7

```
US-09-814-353-20083
; Sequence 20083, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20083
; LENGTH: 9169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-20083
```

```
Alignment Scores:
Pred. No.: 1,45e-12 Length: 9169
Score: 369 Matches: 369
Percent Similarity: 31.20% Conservative: 217
Best Local Similarity: 19.65% Mismatches: 687
Query Match: 3.63% Indels: 606
DB: 12 Gaps: 78
```

```
US-09-617-099b-1 (1-1590) x US-09-814-353-20083 (1-9169)
QY 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
Db 2193 ATGTCAACCCCAAGCTGCCAAGGAGAGACG-GTGAAAAAGGAAACAAAGGTAAAGCCT 2251
QY 21 ProProGlnProGluMetProAsp-----LeuSerHisLeuThrGluGluGluArgLys 38
Db 2252 GAAGACAAAGAGAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGTGGCTTAAAGAGGAGCAAA 2311
QY 39 IleIleLeuAlaValMetAspArgGlnLysLysGluGluGluLysGluGlnSerValLeu 58
Db 2312 ACACCTTATCAAGAGGAGGAGAAACCCAAAAAGGAGAGGTGAAAAAGAAAGTCAAAAAA 2371
QY 59 LysIleLysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThr 78
Db 2372 GAGATCAAGAAAGAGAGAGAAAGAAAGAAAGAAAGAAAGTCAAGAGAAAGTAAAGAGGAGAGAA 2407
QY 79 GluLeuValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGlnLysGlu-ProGln 98
Db 2408 GAGGTTAAGAAAGAAACACCCGCCCAAGGAAGTCAAGAGAAAGTAAAGAGGAGAGAGAA 2467
QY 98 nThrLysLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGln 118
Db 2468 GAAGGAGTGAAGAAAGAGAAAGAAAGAAAGAAAGAAAGTAAAGAGTAAAGAGGAGAGAGAA 2527
QY 118 uSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 138
Db 2528 CCAGAGAAATCATCTACTCTCTGCTGAAGCAAAAAAACCCAGC---TGCTTTAAAC 2584
QY 138 rLysPheAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCys----- 156
Db 2585 CAAGAGTACC-----CAAGAGGAAGAGTCTGTCAAGAAAGATTTCTGTGCTG 2632
QY 157 -----AlaArgCys 159
Db 2633 CCGGAAAGCCAAAGGAGAGGGGAAATAAAGTCAATTAAGAGGAGGCAAGGCCGAG 2692
QY 159 sGlyArgValSerLeuArgSer----- 167
Db 2693 AGGCTGTCTGCTGAGTGTGCGCAGCTGGAGCCACACAGCAGCTGTCTATGGCGGAGCTG 2752
QY 168 -AsnLysValMetTyrValCysAsnLeuCysArgLys-GlnGln----- 181
Db 2753 GAATAGCAGCCATTCGCTGCCCAAGAACTCGAAGCTGAGAGGTCCCTTATGTCTATCTC 2812
QY 182 --GluIleLeuThrLysSerGlyAlaTyrPheTyrAsnSerGlySerAsnThrLeuGlnG 201
Db 2813 CTGAGGATCTAACCAAGGAC-----TTTGAAGAGTTAAAGG 2848
QY 201 InProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLysLysA 221
Db 2849 CTGAAGAGGTGCGATGTAAACAAAGGACATCAAGCTCAGCTGAGCTAATCGAAGACGAAG 2908
QY 221 lalysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerVal----- 238
Db 2909 AGAAACTGAAGGAAACT-----GAGCCAGTCGAGAGCTACGCTCATCCAGA 2953
QY 239 -----ProAlaValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysA 257
Db 2954 AGGAGAGAGAGTCCAAAGGTCTCGGAGTCCCTGATGAGGGAATCATACACATG 3013
QY 257 snGlySerGlyValLysHisGlnIleAlaSerAspMet---ProSerAspArgLysArgS 276
Db 3014 AAGGGGAGGGCGAATGTGAACACAGACACCTGAGGAGCTGGAGCCCTCGAAGAGCGGGAG 3073
QY 276 erProSerValSerArg-----AspGlnAsnArgArgTyrGluGlnSerGluLysArgG 294
Db 3074 TAGACGACATTTGAAAAATTTGAAGATGAAGGAGCGCGTTTTCAGAGAAATCTTCAGAGAC 3133
QY 294 luAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAspTyrA 314
Db -----
```

Db 3134 GAGACTATGAGAG----- 3147
Qy 314 laaspArgSerGlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsnTyrA 334
Db 3148 -----AAGCAGAACTGAGGAGGCTGAGGACCA----- 3177
Qy 334 rgaspSerAsnArgArgGlyHisArgHis-----SerLysGluTyrIleV 349
Db 3178 -----GAAGAGGATGGGAGGAAACACGATATGTGTAGCGCTCCAAAGCAGACAGCCCA 3229
Qy 349 alaspAspGluasp-----ValGluSerArgAspGluTyrGluArgGlnArgArgGluG 367
Db 3230 CTGAGGATGAGAAAGTCCAAAGCGGAGGCTGATGATATCATCAGGAGAGAGGAGT 3289
Qy 367 luGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAlaArgTyrProValLysProG 387
Db 3290 CTGTGGCCAGTGGGATGACCGAGCGGAGGACATGATGATGAGCCATTCGAGAAAGGAG 3349
Qy 387 lnProTyrGluGln-MetArgIleHisAlaGluValSerArgAlaArgHisGluArg 406
Db 3350 AG-----GCTGAACAATCTGAAGCAGGAGGCTGATGAGGAGGACAAAGCTGAAGATGCC 3403
Qy 407 ArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeu--- 425
Db 3404 AGAGAGGAGGAATATGACCCCGGAAAAATGAAGCTGAAGACTATGTGATGGCTGTGTC 3463
Qy 425 ----- 425
Db 3464 GACAAGGCTGACAGGCTGTGTGTCGAGGAGCAGTATGATTCCTCACCACACCAACC 3523
Qy 426 -----LeuArgMetAspArgProSerArgGlnArgSerValSerGluArgArgAlaA 443
Db 3524 AAGCAACTAGGAGGCCACTCTCTGGCCGAGAACCTGCATCTTCAATTCATGATGAGACT 3583
Qy 444 Met-----GluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGln 459
Db 3584 TTACTGAGGCTCAGAGGAGCCACCGCTCTGTGATGAGGAGATCGAGAGACCGAG 3643
Qy 460 GlyGln-----SerTyrProGlnArgThrSerAsnHisSerProPro 474
Db 3644 CTGAGGAATTCACCTGCGACCTCTGGCTACACTCAGTCTACTATTGAGATATCC----- 3697
Qy 475 ThrProArgArgSerProIleProLeuAspArgProAspMetArgArgAlaAspSerLeu 494
Db 3698 -----AGTGACCCACCCCATGGATGAG----- 3721
Qy 495 ArgLysGlnHisLeuAspProSerSerAlaValArgLysThrLysArgGluLysMet 514
Db 3722 -----ATGCTACCCCTCGAGACGTGATG 3745
Qy 515 GluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValArgProPro 534
Db 3746 AGTGATGACCAACATGAAGAGAGCGGAGTCCCTCTCAGGAATTC----- 3793
Qy 535 ProProArgProHisLysSerLysGlyGlyLysMetArgGlnValSerLeuSerSer 554
Db 3794 -----GTAAATATCACCAAA 3808
Qy 555 SerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspAspVal---GluLeu 573
Db 3809 TATGAATCTTCATTGATTCT---CAGGAATACTCTAAACCTGCTGATGTTACACCGCTC 3865
Qy 574 GluSerGluSerValSerGluLysGlyAspSerGlnLysGlyLysArgLysThrSerGlu 593
Db 3866 AACGGATTCTCAGGATCAAAACAGATGCCACTGATGGCAAGGATTACAATGCTTCA 3925
Qy 594 GlnGlyValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysArgMetTyrTyr 613
Db 3926 GCCTCTACCATATCACCCCTCT----- 3949
Qy 614 GlyGlyHisSerLeuGluGluAspLeuGluTyrSerGluProGlnIleLysAspSerGly 633
Db 3950 -----TCCATGGAGGAGAC---AAATTACAGCAGATCTGCTTTACGTGATGCT--- 3994

Qy 634 ValaspThrCysSer-----SerThrThrLeuAsnGluGluHisSerHisSer 649
Db 3995 -----TACTGCTCTGAAGTGAAGCAGCACCTTTGGACATCAAGATAGCATCTCA 4048
Qy 650 AspLysHisProValThrTrpGlnProSerLysAspGlyAspArgLeuIleGlyArgIle 669
Db 4049 GCTGTTTCAAGTGAAGAGTCCAGCCATCGAAG----- 4081
Qy 670 LeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSerGlyAlaMetLeuGly 689
Db 4082 -----AGCCGCTCCTGAGTCCATCTCCACCATCACCTTAGAAAAAGACCCCTGGGT 4135
Qy 690 LeuLysValValGlyGlyLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLys 709
Db 4136 GAACGTAGTGTGAACCTCTCTGACGCCCAAT----- 4168
Qy 710 ValLysLysGlySerLeuAlaAspThrValGlyHisLeuArgProGlyAspGluValLeu 729
Db 4168 ----- 4168
Qy 730 GluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleLeu 749
Db 4168 ----- 4168
Qy 750 GluSerLysProGluProGlnValGluLeuValValSerArgProIleGlyAspIlePro 769
Db 4169 -----GAGATTAAAGTCTCTCGAGGCGAGAGTAGCCCGC 4204
Qy 770 ArgIleProAspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGln 789
Db 4205 GTGCTCTCTGAGGTGACCCAGAGTAGTGTGAACAATTGTGTACTCTCTGAGACAG 4264
Qy 790 LysMetAsp-----ArgProSerIleSerValThrSerProMetSerProGlyMetLeu 807
Db 4265 ACTCTGGAAGTGGTGCACCATCTCAGTCCGTGACT-----GGCAGTGCT 4309
Qy 808 ArgAspValProGlnPheLeuSerGlyGlnLeuSerIleLysLeuTrpPheAspLysVal 827
Db 4310 GGTCAACACCTTACTATCAATCTCTACT-----GACGAGAAATCC 4351
Qy 828 GlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArgGluAspGly 847
Db 4352 AGTCAT-----CTCCCTACAGAGTCATTGAA 4378
Qy 848 ArgProArgAsnProTyrValLysIleTyrPhe---LeuProAspArgSerAspLysAsn 866
Db 4379 AAACCA-----CCAGCAGTTCAGTGATTTTGAATTGATGATGATGATGATGATGAT 4432
Qy 867 LysArgArgThr-----LysThrValLys 874
Db 4433 GAAAGGGCTTCAGTAAGCCCATGATGAGCCCGTCTGACTCAGAGTCTCTTATTGAA 4492
Qy 875 LysThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGlu 894
Db 4493 AAAGTTTGTCTCTTTACGAGCCCGCCCTCATTTGATCCGAGTCTGCTATGAAAGT 4552
Qy 895 PheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluGluGlu 914
Db 4553 TTT-----CTAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4594
Qy 915 SerGluPhe-----LeuGlyGluIleLeu 922
Db 4595 AGTCTTTTGAAGAAAGAGTGGAAAAACAAGGCTCTCCAGACCAAGTAAGTCCAGTTCT 4654
Qy 923 IleGluLeuGluThrAlaLeuLeuAspAsp----- 932
Db 4655 GAATGACTTCTACTAGTCTTTTACAGACAAACAGGAGGAAAGACACAGACTTTTCA 4714
Qy 933 -----GluProHisTrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeu 949
Db 4715 CCATAAAGAGAGACTTTGGCCCAAGAAAAAAGAACTGATGATGATGATGATGATGATGAT 4771

Qy 950 ProArgProSerProTyrLeuProArgArgGlnLeuHisGlyGluSerProThrArg--- 968
Db : : : : :
4772 TCTCAACAGCACTGGCTCTGATGAAGAAATAGGATGTTTCTCCACAAATA 4831
Qy 969 -----ArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db : : : : :
4832 GATGTCAGTCAGTTGGATCTTTAAAGAGACACTAAGATGTCCATTTCTGAAGGTACT 4891
Qy 981 ValSerAspTyrAspCysGluAsp---GlyValGlyValValSerAsp---TyrArgHis 998
Db : : : : :
4892 GTCTCAGCAAGTCACTGCTCTCTGTTGATGAGGCGTAGCAGAGACACGATCTCTCAT 4951
Qy 999 AsnGlyArgAspLeuGlnSerSerThrLeuSerVal-----ProGluGln 1013
Db : : : : :
4952 ATGAGAGGTGTGGCTCAGTCTCCACAGCCTCAGTGGCTACGAGTCAATTTCCAGAGCCA 5011
Qy 1014 ValMetSerSerAsnHisCysSerProSer-----GlySerProHisArg 1028
Db : : : : :
5012 -----ACAACAGATGATGTCTCTCCATCTCTGCATCTGCTGAGTTTGGCTCCCAATTC 5065
Qy 1029 ValAspVal-----IleGlyArgThr 1035
Db : : : : :
5066 ACAGAGTAGATGACTCCCTTTCAGTGTCTGTGTGMAAACACCTACCACATTCAGGAA 5125
Qy 1036 ArgSerTrpSerProSer-----AlaProPro--- 1044
Db : : : : :
5126 ACAGAAATGTCTCATCTAAAGAGAAATGCCAAGACCGATGTCAATTTCTCCACAGAT 5185
Qy 1045 -----ProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAla 1058
Db : : : : :
5186 TTCTCCCTAAACATGCAAGTCCAGGACACCGTTCAAGATCACAGATCTGAACAGTCC 5245
Qy 1059 ThrGlyHisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyr 1078
Db : : : : :
5246 TCA-----ATGCTATTATGAATTT 5263
Qy 1079 SerSerAspArgAspCysGluAlaAlaAsp-----ArgGlnProTyrHis 1095
Db : : : : :
5264 GGCCAAAGATCTCTGACATCCCTGCTATGACTTCAGTCAGCAGTCTCCAGATCAC 5323
Qy 1096 ArgSer-----ArgSerThrGluGlnArgProLeuLeuGluArgThr 1109
Db : : : : :
5324 CCTACAGTGGGTGCAGGCGTCTTCACATCATCTGAAATGGCCA-----ACT 5371
Qy 1110 ThrThrArgSerArgSerSerGluArgProAspThrAsnLeuMetArgSerMetPro--- 1128
Db : : : : :
5372 GAAGTGGACTACAGTCTCTCTGACATGCAGGACTCCAGTTTATCACATAAGATACCACCT 5431
Qy 1129 -----SerLeuMetThrGlyArgSer 1135
Db : : : : :
5432 ATGAGAGCGCTCTACACCCAGATATATGATCTTCTGAGCTCATCTCAGTATCTCAG 5491
Qy 1136 AlaProProSerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSer 1155
Db : : : : :
5492 GTAGAGGCTCCCGTCCACCTCTCTGCTCAT-----ACCCTTCTCAGATCGCTTCT 5545
Qy 1156 ProSerSerThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLys 1175
Db : : : : :
5546 CCT----- 5548
Qy 1176 GlyThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeu--- 1194
Db : : : : :
5549 -----CTCCAAGAGATACTCTATCCGATGTGTCTCTCCAGAGATATCTCTTATAT 5602
Qy 1195 -----AsnLysTyrLysGlnValAlaGlySerAspProArgLeuGluGln 1209
Db : : : : :
5603 GCCTCACTCACCTCTGAAAAGTGCAGAGTCTGGAAGGAGAGAGCTCTCTCCAAATCT 5662
Qy 1210 AspTyrHisSerLysTyrArgSerGlyTrpAspProHisArgGlyAlaAspThrValSer 1229
Db : : : : :
5663 GATATC-----TCTCCACTCACCACGAGAGTCTCTCTCTTATATATCA 5707
Qy 1230 ThrLysSerSerAspSerAspValSerAspValSerAlaValSerArgThrSerSerAla 1249

Db 5708 CCTACTTTTTCAGATTCT-----ACCTCTGCACTCAAGAGAGAAACAGCAACT 5755
Qy 1250 SerArgPheSerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArg 1269
Db : : : : :
5756 TGCACAGATTCTCTCTCCACCAATAGATGCAGCATCCGACAGGCCCTATGGCTTCCGT 5815
Qy 1270 LysIleSerValPheThrSerLysMetGln-----AsnArgGlnMetGly 1284
Db : : : : :
5816 ---GCCTCAGTGTATTTCGATACATGCACACCATCTAGCCTTGAATAGAGATTGTTC 5872
Qy 1285 ValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMetCys----- 1301
Db : : : : :
5873 ACACCTGGCTCGAGAGGACAGTGGAGGAGACACACCTGCTGACTTTGATGATGCCTAT 5932
Qy 1302 ---SerLeuGlnLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320
Db : : : : :
5933 CAAAAGCCTCAGGAAACACACAGTCCCGAGATGAAGAAGATTATGACTATGACTTAT 5992
Qy 1321 SerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
Db : : : : :
5993 GAGNAGACACCCGACCTCAGATGGTGGTGGCTATTACTATGAGAAGATAGAGAAACC 6052
Qy 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArg 1360
Db : : : : :
6053 ACAAATCTCCAGTGCAGTGGCTACTCTAT----- 6085
Qy 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
Db : : : : :
6086 GAGACCATTTGGAAATACAGACCCCTGAAGATGGTGACTATTCTTATGAAATTTAT 6145
Qy 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn 1400
Db : : : : :
6146 GAGAGACACCGACCCCTGGAAGGGTGGTACTCATATGACATAAGTGAAGAACACC 6205
Qy 1401 LeuIlePheProGlyVal-----ArgLeu 1408
Db : : : : :
6206 ACCAGCCCCCGAAGTGTGTTACAGCTATGAAAAGACTGAGAGGTCTAGAGGCTT 6255
Qy 1409 AlaSerAsp-----SerGlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeu 1426
Db : : : : :
6266 CTGATGACATCAGCAATGGCTATGATGACTCTGAGGATGCT----- 6307
Qy 1427 ValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAsp 1446
Db : : : : :
6308 ---GGCCACACACACTTTGGGACCCCGACTACT-----CTTATGAA 6346
Qy 1447 LysLysGlyGlnLeuGluValIleIleArgAlaArgGlyLeuValValLysProGly 1466
Db : : : : :
6347 CCCTGAGAAATACAGTTCCCTG-----AGT 6376
Qy 1467 SerLysThrLeuProAlaProTyrVal-----LysValTyrLeuLeuAspAsn 1482
Db : : : : :
6377 CTGAAGGTATTCTATGAGACATCTACAAAGACACCAACCCCTGATATCTCCACAT 6436
Qy 1483 GlyValCysIleAlaLysLysThrLysValalaArgLysThrLeuGluProLeu--- 1501
Db : : : : :
6437 ACTGTT-----ACGAGACTGCAGAGAAATCCTAGAACCCCTCAGG 6478
Qy 1502 -----TyrGlnGlnLeuLeuSerPheGluGluSer 1511
Db : : : : :
6479 CATCCACATTTCTACGACCTTCAGACCTATGCTACTGTCAG-----AAAGAAGT 6532
Qy 1512 ProGlnGlyArgValLeuGlnIleValTrpLysAspTyrGlyArgMetAsp----- 1529
Db : : : : :
6533 CCCCTCAGAAGCCCGTCAGAGTGCATTTATGCTCTCGTCTCTCTTGTGATACAGC 6592
Qy 1530 -----HisLysSerPhe-----MetGlyValalaGlnIleLeuLeuAspGluLeu--- 1544
Db : : : : :
6593 ACCCAAGACAGAGCTTTCCCTCTTTTCAATATCCCAATCTCTCTTGAAGTGGTTGCCA 6652
Qy 1544 ----- 1544

Qy 356 rArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGln-----AlaAr 372
Db 3804910 GCGGTCTCGCTCATACAGCGCCCGCGAGCCCGCGATGCTCGGTGATCATCAGGAGGAG 3804969
Qy 372 gTyrArgSerAspProMetLeuAlaArg-TyrProValLysProGlnProTyrGlu----- 390
Db 3804970 GCGACGTTCCGACAGCGTCCGCGCATGATCATCGGAACCGCGCGCATGACCGT 3805029
Qy 391 -----GluGlnMetArgIleHisAlaGluValSerArgAlaArgHisGluArg- 406
Db 3805030 GCCCGCGGACGAATTCGCGTCTCTTCAGGACAGTGTATCGACCGGACCAACCGCGCG 3805089
Qy 407 -----ArgHisSerAspValSerLeuAla----- 414
Db 3805090 GTGATCATCCACCGAGCGCGAGCAGTCCCGCGAGGACGACCCAGCCCGCTGGG 3805149
Qy 415 -----AsnAlaGluLeuGluAspSerArgIleSerLeuL 426
Db 3805150 ATCCGCGGTGAGACGCGCGGATGATGAGGAACGACGCGGAGACGACGACCA-AGTCGAC 3805208
Qy 426 euArgMetAspArgProSerArgGln-----ArgSerValSerGluArgAlaAlaM 444
Db 3805209 GAAGAACATCAGCCCGCGTCTCGTCAGCCCGCGGAGTACCCCGACAGCGCGCGAGAA 3805268
Qy 444 etGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSert 464
Db 3805269 TCGCGGAGCCCGTCTCGGAGAACCGCGCACCAACGACCGCGATCAGCAACGACTTCTGAG 3805328
Qy 464 yrProGlnArgThr-----SerAsnHisSerProProThrProA 477
Db 3805329 TCCCGCAGCGTCTCGCGAACACGCTCTGCGCGATCGGTGTCCTCCACCAAGTGTCT 3805388
Qy 477 rArgSer-ProIlePro-----LeuAspArgProAsp 487
Db 3805389 GCGCTCGCGCCCGACCGACCGACGCGGTGTAGTCGATAGGTGTACGACCAACCGGGGTAC 3805448
Qy 488 MetArgArgAlaAspSerLeuArgLysGlnHis-----LeuAspProSerSerAl 505
Db 3805449 GTACGCGCCGACGAGCGACGAGAACAGCAGTACGAGGACGAGCGATTCGCGGACGCG 3805508
Qy 505 aValArg-----LysThrLysArgGluLys-MetGluThrMetLeuArgAsnAspSerLeuS 524
Db 3805509 GCGCGGATTCGTACGAAGCGCGCGACGACGCGTCCCGCGCGCGACCGGAC----- 3805561
Qy 524 erSerAspGlnSerGluSerValArgProProProProArgProHisLysSerLysG 544
Db 3805562 -----GTACGCGCGCTCGACCG----- 3805579
Qy 544 lyGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAlaSerThrProG 564
Db 3805580 -----TTCGACCAAC----- 3805589
Qy 564 luTyrThrSer-----CysAspAspValGluLeuGluSer-----GluSerValS 579
Db 3805590 AGGGGTGATCGTCATGACGCTCGCGGATCGCGGTCGAGCGCGGCGTGCAGCACGCTCC 3805649
Qy 579 erGluLysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGly----- 595
Db 3805650 GCGAGGAACCGCGACACGACGCGTACCGCGCGAAGAGATTGACCGGCGGACGACCGGAG 3805709
Qy 596 -----ValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysMetTyrT 613
Db 3805710 TTGACGTCGTTCTTCGCGATCGA----- 3805732
Qy 613 yrGlyGlyHisSerLeuGluGluAspLeuGluTyrSerGluProGlnIleLysAspSerG 633
Db 3805733 --GGAGATGAACACTCTCCCATCCCGTGCAC-----GCCGAAGATCGTCTCCGTGAGATG 3805787
Qy 633 lyValAspThrCysSerSerThrThrLeuAsnGluGluHisSerHisSerAsp-----LysH 652
Db 3805788 GCACCGGTGAAGAGCGCGAGAACCGCTACGAGAAGTACGTCGACATGGGGATCAGCGCC 3805847

Qy 652 isProValThrTrpGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuA 672
Db 3805848 GTCOG-----CACCGCGTCTCAGCAGCACCCCGCGCGCGCGAGACCCCTT----- 3805894
Qy 672 snLysArgLeuLysAspGlySerValProArgAspSerGlyAlaMetLeuGlyLeuLysV 692
Db 3805895 -----GGCCCGCGC-----GGTCCGCGAGA 3805913
Qy 692 alValGlyGlyLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLysValLysL 712
Db 3805914 TAGTCGAGCCCGACACGTCGACGATGTT-----G 3805943
Qy 712 ysGlySerLeuAlaAspThrValGlyHisLeuArgProGlyAspGluValLeuGluTrpA 732
Db 3805944 CTCGCTGTGTAGCGCTGTACCGCGCGATCGCGAAGCGCGACGGA----- 3805990
Qy 732 snGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGluSerL 752
Db 3805991 -----CAGGCTGGCGAG-----CAGCAGGTGTACCGCGCGCTCTCGAGTACG 3806033
Qy 752 ysProGluPro-----GlnValGluLeuValValSerArgProIleGlyAspIle----- 768
Db 3806034 GCCCGGAACCGCGCGCTCTCTCGCGCGGCTCTCTCGCGCGTGAACCTGATCAGATCCGTC 3806093
Qy 769 -----ProArgIleProAspSerThr-----HisAlaGlnLeuGluS 781
Db 3806094 CCTGTCTCGGTGTACACAGATCGTCCGCTTGTGAAGAGATCGCCACGACGAAACACG 3806153
Qy 781 erSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSerP 801
Db 3806154 GCGGTGTGAGACGAGTGTGAAGAGCGGACGTCACCGTCCGCTCGGTGAGCGGTACTGC 3806213
Qy 801 ro-----MetSerProGlyMet-----LeuArgAsp-ValProGlnPheLeuSerGly 816
Db 3806214 CGCACCGCGCTCCAGGCTCCGATCAGCACCGCGCGACCGTTCGCGACGCTTCGCGCG 3806273
Qy 817 GlnLeuSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeu 836
Db 3806274 AGCAGCAGCGCGAGCTG----- 3806291
Qy 837 GlyAlaLysAspLeuProSerArgGluAspGlyArgProArgAsn-----ProTyrVal 854
Db 3806292 -----ACCCCGATGCG-----CGCCCGAAGTCTCCGCTTGCCTGCTGTA 3806330
Qy 855 LysIleTyrPheLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLys 874
Db 3806331 CCGTCGATCGTCTCGCG-----AGGTCTCGCGCGACCGCATCCCGCGCGACGAG 3806387
Qy 875 LysThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProVal-----HisArg 892
Db 3806388 CGCGCGACGACGACCGCGTGTCTGTGACGCGATCTC-GGTGAGTGTGTGATTCACCGC 3806446
Qy 893 ArgGluPheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGlu 912
Db 3806447 GCTCGCGGAGACGCGCGCTCGCGCGCTCGTAGTA-----CGCGCGGGTTCGAGCGC 3806500
Qy 913 GluGluSerGluPheLeuGlyGluIleLeuLeuGluThrAlaLeuLeuAspAsp 932
Db 3806501 GAACGAGGCGCAGCATAGAGAGGACGCGCGCGCGACGAGCAGCAGCGCGTAATAGCC 3806560
Qy 933 GluProHisTrpTyrLysLeuGlnThr----- 941
Db 3806561 GAGCCGCT-TGGCCAGATAGGCGAGCGACGTCATCGCTCCCGTCCCGTGTTCACCA 3806619
Qy 942 -----HisAspValSerSerLeuProLeuProArg---ProSerProTyrLeu 956
Db 3806620 GCATGAAGAGACATACAGCAGTCCAGAACCGCTCCACGACTCCCGTAGCGGTCTTAC 3806679
Qy 957 ProArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIle 976
Db 3806680 GTTCGCGCG-----ATTACGCGATGACCCACG----- 3806706
Qy 977 SerAspSerGluValSerAspTyrAspCysGluAspGlyValGlyValSerAspTyr 996

Db 3806706 ----- 3806706
QY 997 ArgHisAsnGlyArgAspLeuGlnSerThrLeuSerValProGluGlnValMetSer 1016
Db 3806706 ----- 3806706
QY 1017 SerAsnHisCysSerProSerGly-----SerProHisArgValAspValIleGly 1033
Db 3806707 -----GCTCCGCGCCGCCACCGAGAGACACCCACAAA-----GGA 3806742
QY 1034 ArgThrArgSerTrpSerProSerAlaProProGlnArgAsnValGluGlnGlyHis 1053
Db 3806743 CACCGACATGCCGAGAACCCCGCGCGACCCCTGGTGGCGGC-----CAC 3806787
QY 1054 ArgGlyThrArgAlaThrGlyHisTyrAsnThrIleSerArgMetAspArgHisArgVal 1073
Db 3806788 CGCGGC-----CACCGCGTT 3806802
QY 1074 MetAspAspHisTyrSerSerAspArgAspArgAspCysGluAlaAspArgGlnPro 1093
Db 3806803 GCTGTGTACCGGCTCGGTTCTCGACGACCAAGCGCGAGCGCGAGA-AGAGCGCG 3806861
QY 1094 TyrHisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThr---ThrArg 1112
Db 3806862 -----CCCCGCGCGGAGCCGACGATCAACGCCACCCGCG 3806897
QY 1113 SerArgSerSerGluArgProAspThrAsnLeuMetArgSerMetProSerLeuMetThr 1132
Db 3806898 TCAGCGACATCCAGCAGCGCGCCACCTCAAGTTCGGATCTCCCGATACATCGCGACT 3806957
QY 1133 GlyArgSerAlaProProSerProAlaLeuSerArgSerHisProArgThrGlySerVal 1152
Db 3806958 TCAACCGCTACACCGCCACG-----GCAACCGGGTGACCGCGCGATCAACAGAGC 3807011
QY 1153 GlnThrSerProSerSerThrProGlyThr-GlyArgArgGly-----ArgGlnIle 1169
Db 3807012 TGGTCAGGCCACAGCTTTCACCGCGCGCGCGCGCGCTCCGCGCGAAGAGAACT 3807071
QY 1169 uProGlnLeu-----ProProLysGlyThrLeuGluArgSerAl 1182
Db 3807072 TCCTGCTCCCGCTCGGTGACCTCCACCAGCCCGCCAGGTCTCACTTCAAGCTCAATC 3807131
QY 1182 aMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaG 1202
Db 3807132 CCAAGGCCAAGTGTCCGACGCGCGCCCTGAGCTGGCGGACCTCCGTCGCTGTGA 3807191
QY 1202 ySerAspProArgLeuGluGln-AspTyrHisSerLysTyrArgSerGlyTrpAspPro- 1221
Db 3807192 AGGCCACGAAACGCGACGAAGAGTACGAGCGCGCCCAACACCTCCGCTACGACCGA 3807251
QY 1222 -----HisArgGlyAlaAspThrValSerThrLys---SerSerAspSerAspV 1237
Db 3807252 TCAGCGCGTCCGCGAGCGGACCGACCGCCACCGTCAAGTCACTTCTCTCGCGT 3807311
QY 1237 alSerAspValSerAlaValSerArgThrSerSerAlaSerArgPheSerThrSerT 1257
Db 3807312 ACGCCACTG-CAAAGCGGCTTTCGACCCACTGCTCCCGCGCGG-GGATCGACACCC 3807369
QY 1257 yrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerL 1277
Db 3807370 GGGCAAGTTCAACAAGGGTGGACCGAGAGATCCCG-----TCACCGG 3807414
QY 1277 yMetGlnAsnArgGlnMetGlyValSerGly-----LysAsnLeuThrLysSerThrS 1295
Db 3807415 CGCGCGCTTCAGATCTCCGCTACGACAGAGCGCGCCAGACCATCACGGCGGTCCCGA 3807474
QY 1295 erIleSerGlyAspMet-----CysSerLeuGluLysAsnAspGlySerGlnSerAspT 1313
Db 3807475 CCCCCAGTGTGGGGTACGAAGCCCAAGCTGGACTCGTCTTCGCGCGCTGGACTT 3807534
QY 1313 hrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSerSerIleGlyAlaLysM 1333
Db 3807534 GACGAGTTCCGACGAGG 3808561
Db 3807535 CAGCGCTGGAGCGGACGCTACCTCAACAAGAGATCGACTACGCTCCGCTCCGCTCC 3807594
QY 1333 etValAlaIleValGlyLeuSerArgLys----- 1342
Db 3807595 CGAGGACTCAAGCGCTTGCACAGCGCGCGGACGACATCGCACCGCGCGCGCGGTG 3807654
QY 1343 --SerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyGlyLys-----LysLeuA 1360
Db 3807655 GGACGAGTGCATATCAC-CTGAACGCGCGCGCGCGCTGAAGACGCTCCGGGTCC 3807713
QY 1360 rgSerThrValGlnArgSerThrGlu-----ThrGlyL 1371
Db 3807714 GCCAAGCGTCCACGACGCGCATCGACGCGCATCGCCCGCTTCGGCAAGAC 3807773
QY 1371 euAlaValGluMetArgAsnTrp-----MetThrArgGlnAlaSerArg 1386
Db 3807774 TCCCGTGCAGCTGAAGACGCTACGCGACGCTTCTTCTATCCCAACGACCGCGCTAC 3807833
QY 1386 luSerThrAspGlySerMetAsnSerTyrSerSerGlu-----GlyAsnLeuIle--P 1403
Db 3807834 AGGACACTCCGGTACGTACGCGACGCTACGACGCTCGAGCGCGCGCGAACTCTCTCGACG 3807893
QY 1403 heProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGlyP 1423
Db 3807894 CGCGCGGTGGAAGACCAAGCGCGACGCGCGGTGAAGAC----- 3807934
QY 1423 roAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValG 1443
Db 3807935 -----GGCAAGCAGCTCGCCCTCACCTACGCTACGCGCGCGCGACACT 3807980
QY 1443 lyMetMetAspLysGlyGlnLeuGluValGluIleAlaArgAlaArgGlyLeuValV 1463
Db 3807981 CCGCGAGAGACGAGCGCGAGCTGTGTGCGAGCGGCGAGCTCGCGCGCTCGGG-----A 3808034
QY 1463 allysProGlySerLysThrLeuPro-----AlaprotYrValLysValT 1478
Db 3808035 TCAAGTGTCCATCAAGAGTGTGCGACGAGAGCTACTTCTTCAAGTACGTCACTACG 3808094
QY 1478 yrLeuLeuAspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrL 1498
Db 3808095 GCAACTTCAC-----CTGTGCTGCTGCGTGCAGACAGATCTTCCGCTCCGAGG 3808148
QY 1498 euGluProLeuTyrGlnGln----- 1504
Db 3808149 CGTACCGGATCTACCGCGACCGCGCAAGAACCTGTACGAGAACTACGCTCGGTG 3808208
QY 1505 -----LeuLeu-----SerPheGluGluSerProG 1513
Db 3808209 GCTCCCGGAGATCGACACCTGCTGAAGAGCGGGTGAAGACCGACCGACCGCGCGAG 3808268
QY 1513 lnglyArgValLeuGlnIleIleValTrpGlyAspTyrGlyArgMetAspHisLysSerP 1533
Db 3808269 AGGCCAAGATC-----TACAAGAGCGCGCGACAGAGATCT 3808304
QY 1533 he---MetGlyValAla-----GlnIleLeuLeuAspGluLeuG 1545
Db 3808305 GGGCCCTCGGCGACCTCATCCGCTCTACCGCGCGCGCTTCCGCTCCGCTCGA 3808364
QY 1545 luLeuSerAsn-----MetValIleGlyLysP 1554
Db 3808365 ACCTCGCCACTACCGCGCGGGGCGCTCGCGCGACGACGCTTACCAAGGTGCGTGGC 3808424
QY 1554 heLysLeuPheProPro-----SerSerLeuValAspP 1565
Db 3808425 TCAAGAAGAGTGAACGCTCCGCGACACCCAGCGGTGGAGCGGGTTCAGCGCGCTTCAC 3808484
QY 1565 roThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGluSerSerThrGlyP 1585
Db 3808485 CAGGTCCGGGACCCCGCTCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAC 3808544
QY 1585 roSerTyrSerArgSer 1590
Db 3808545 GACGAGTTCCGACGAGG 3808561

RESULT 9
 US-10-133-937-37
 ; Sequence 37, Application US/10133937
 ; Publication No. US20030207278A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Khan, Javed
 ; APPLICANT: Ringner, Markus
 ; APPLICANT: Peterson, Carsten
 ; APPLICANT: Meltzer, Paul
 ; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
 ; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
 ; TITLE OF INVENTION: OTHER BIOLOGICAL STATES
 ; FILE REFERENCE: 11613.56US01
 ; CURRENT APPLICATION NUMBER: US/10/133,937
 ; CURRENT FILING DATE: 2002-11-04
 ; NUMBER OF SEQ ID NOS: 99
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 37
 ; LENGTH: 9161
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-133-937-37

Alignment Scores:

Pred. No.:	3,78e-12	Length:	9161
Score:	295.00	Matches:	366
Percent Similarity:	31.37%	Conservative:	221
Best Local Similarity:	19.56%	Mismatches:	691
Query Match:	3.56%	Indels:	595
DB:	12	Gaps:	78

US-09-617-099b-1 (1-1590) x US-10-133-937-37 (1-9161)

QY	1	MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro	20
DB	1847	ATGTCAACCAACCAAGCTGCCAAGGAGAGACG-GTCAAAAAGGAAACAAAGGTAAAGCCT	1905
QY	21	ProProGlnProGluMetProAsp-----LeuSerHisLeuThrGluGluArgGly	38
DB	1906	GAAGACAGAAAGAGAGAAAGAAAGCCAAAGAAAGTGGCTTAAAGGAGGACAAA	1965
QY	39	IleIleLeuAlaValMetAspArgGlnLysGluGluGluLysGluGlnSerValLeu	58
DB	1966	ACACCTATCAAGAGAGAGGAAACCAAAAGAGAGAGGTGAAAAAGAGTCAAAAA	2025
QY	59	LysIleLysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThr	78
DB	2026	GAGATCAAGAAAGAGAGAAAGAAAGAACCCCAAGAA-----	2061
QY	79	GluLeuValAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGln	98
DB	2062	GAGTTAAGAAAGAAACACCGCCAAAGGAGAGTCAAGAGGAAGTTAAGAGGAGAGAG	2121
QY	99	ThrLysLeuHisGlnGlnPheGluMetTyLysGluGlnValLysLysMetGlyGluGlu	118
DB	2122	AAGGAAGTGAAGAGGAAAGAGAACCCAAAGAAAGAAATTAAGAGTCCCTTAAGAC	2181
QY	119	SerGlnGlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThr	138
DB	2182	GCAAGAAATCATCTACTCTCTCTCTGTAAGCAAA-AAAACACAGC--TGCTTTAAACC	2237
QY	139	LysPheAlaaspGlyCysGlyHisAsnCysSerTyCysGlnThrLysPheCys	156
DB	2238	AAAAGTACC-----CAAGAGGAGAGTCTGTCAAGAAAGATCTGTCTCTG	2285
QY	157	-----	159
DB	2286	CGAAAGCCNAAGAGAGGAGGAAATTAAGTATTAAAGAGAGAGGAGCCCGCAG	2345
QY	160	GlyGlyArgValSerLeuArgSer-----	167
DB	2346	GGCTGTGCTGACGCTCGGCACTGGAGCCACCAGCAGCTGTCTATGGCGGCACTGG	2405

QY	168	AsnLysValMetTrpValCysAsnLeuCysArgLys-GlnGln-----	181
DB	2406	AATAGCAGCCATTGGCCCTGCCAAAGACTCGAAGCTGAGAGTCCCTTATGTCATCTCC	2465
QY	182	-GluIleLeuThrLysSerGlyAlaTrpPheTyAsnSerGlySerAsnThrLeuGlnG	201
DB	2466	TGAGGATCTAACCAAGGAC-----TTTGAAGAGTTAAAGGC	2501
QY	201	nProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGlnLysAl	221
DB	2502	TGAAGAGTTCGATGTAAACAAAGACATCAAGCCTCAGCTGGAGCTAATCCAGACGA	2561
QY	221	atLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerVal-----	238
DB	2562	GAAACTGAAGGAAACT-----GAGCCAGTCGAAGCCTACGTCATCCAGAA	2606
QY	239	-----ProAlaValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLys	257
DB	2607	GGAGAGAGAGTCAACCAAGGCTCTCCGAGTCCCTGATGAGGAGATCACTACCACTGA	2666
QY	257	ngLysSerGlyValLysHisGlnIleAlaSerAspMet-----ProSerAspArgLys	276
DB	2667	AGGGAGGCGCAATGTGAACACAGACCTGAGGAGCTGGAGCCCTCGAGAAGCAGGAGT	2726
QY	276	rProSerValSerArg-----AspGlnAsnArgArgTyGluGlnSerGluGluArgG	294
DB	2727	AGACGACATGTAAATTTGAAGATGAAGAGCGCGGTTTGAAGAAATCTTCAGAGACTGG	2786
QY	294	uAspTySerGlnTyValProSerAsp-----	303
DB	2787	AGACTATGAAGAGAGGAGGAGAACTGAGGAGGCTGAGGAGCCAGAGAGATCGGAGGA	2846
QY	304	-----GlyThrMetProArgSerProSerAspTyAlaAspArgArgSerG	319
DB	2847	ACAGTATGTGTGAGCGCCTCCAAGCACAGCCCACTGAG--GATGAGAAAGTCCCAA	2903
QY	319	nArgGluProGlnPheTyGluGluProGlyHisLeuAsnTyArgAspSerAsnArg	339
DB	2904	GCGGAGGCTGATGCATACATCAGGAGAGAGGAGTCTGTGCGCAGTGGGATGACCG	2963
QY	339	gGlyHisArgHisSerLysGluTyIleValAspAspGluAspValGluSerArgAspG	359
DB	2964	AGCCGAAGAGACATGATGAGGCCATTCGAAAGAGAGAGGCTGAACAAATCTGAAGAG	3023
QY	359	uTyGluArgGln-----ArgArgGluGluGluTyGlnAlaArgTyArg	374
DB	3024	GGCTCATGAGGAGGACAAAGCTGAAGATGCCAGAGAGGAGGAATATGAG-----	3072
QY	374	gSerAspProAsnLeuAlaArgTyProValLysProGlnProTyGluGlnGlnMet	394
DB	3073	-----COGAAATAATGAAGCTGAAGACTA	3098
QY	394	gIleHisAlaGluValSerArgAlaArgHis-----GluArgArgHisSerAs	410
DB	3099	TGTGATGCTGTGTGTCACAGAGCTGTCAGAGGCTGTGTGTCGAGAGCTGATGATT	3158
QY	410	pValSerLeuAlaAsnAlaGluLeuAspSerArgIleSerLeuLeuArgMetAsp	430
DB	3159	CCTCACCACCAACCAAGCAA-----CTAGGAGCCAGTTC	3194
QY	430	gProSerArgGlnArgSerValSerGluArgArgAlaAlaMet-----GluAs	446
DB	3195	TCCTGGCCGAGAACCTGTCATCTTCAATTATCATGAGACTTTACCTGGAGGCTCAGAG	3254
QY	446	ngLysSerTySerMetGluArgThrArgGluAlaGlnGlyGln-----	461
DB	3255	CGAGGCCCGCTTCTGTATGAGGAGAAATGAGAGACCGCTGAGGAATTCAGTCCAC	3314
QY	462	-SerSerTyProGlnArgThrSerAsnHisSerProThrProArgArgSerProI	481
DB	3315	CTCTGGTACACTGACTTACTATTGAGATATCC-----AGTGAGCCAC	3359

Db 4995 GCTTCACATCACTGAAATGGGCA-----ACTGAAGTGAAGTACAGTCTCCTC 5042
 Qy 1116 rGluArgProAspThrAsnLeuMetArgSerMetPro----- 1128
 Db 5043 TGACATGACGAGCTCCAGTTTATACATAGATACACCTATGAGGAGCGGTCTACAC 5102
 Qy 1129 -----SerLeuMetThrGlyArgSerAlaProProSerProAlaLe 1142
 Db 5103 CCAAGATAATGATCTTTCTGAGCTCATCTCAGTATCTCAGTAGAGGCTCCCGGTCCAC 5162
 Qy 1142 uSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrProGlyTh 1162
 Db 5163 CTCTCTGCTCAT-----ACCCCTCTCAGATCGCTTCTCCT----- 5199
 Qy 1162 rGlyArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArgSerAl 1182
 Db 5200 -----CTCCAGAGATAC 5213
 Qy 1182 aMetAspIleGluGluArgAsnArgGlnMetLysLeu-----AsnLy 1196
 Db 5214 TCTATCCGATGTTGCTCTCCACAGATATGTCCTTATATGCTCACTCACTCACTCACTCA 5273
 Qy 1196 sTyLysGlnValAlaGlySerAspProArgLeuGluGlnAspTyHisSerLysTyAr 1216
 Db 5274 AGTCAAAGTCTGGAAGGAGAGAGCTCTCTCCAAAATCTGATATC----- 5319
 Qy 1216 gSerGlyTrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAs 1236
 Db 5320 -TCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 5376
 Qy 1236 pValSerAspValSerAlaValSerArgThrSerSerAlaSerArgPheSerThrSe 1256
 Db 5377 -----ACCTCTGCACTCAAGAGAGAAACAGCAACTTGCACAGTTCTCTCTCTCC 5426
 Qy 1256 rTyMetSerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSe 1276
 Db 5427 ACCAATAGATGACATCCGACAGGCCATGTCCTCCGT---GCCTCAGTGTATTGCA 5483
 Qy 1276 rLysMetGln-----AsnArgGlnMetGlyValSerGlyLysAsnLeuth 1291
 Db 5484 TACAATGCAACACCATCTAGCTTGAATAGAGATTGTCCACACCTGGCCCTGGAAGAGA 5543
 Qy 1291 rLysSerThrSerIleSerGlyAspMetCys-----SerLeuGluLysAsnAe 1307
 Db 5544 CAGTGGAGGAGACACACCTGTGACITTAGCTATGCTATCAAAAGCCTGAGAAACACAC 5603
 Qy 1307 pGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysArgArgSe 1327
 Db 5604 CAGGTCCTCCAGATGAAGAGAGATTATGACTATGAGTCTTATGAGAGACACACCGGACCTC 5663
 Qy 1327 rSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlase 1347
 Db 5664 AGATGTGGTGGCTATTACTATGAGAGATAGAGAGAACCAAAATCTCCAAGTACAG 5723
 Qy 1347 rGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerth 1367
 Db 5724 TGCGTACTCTAT-----GAGACCATTTGGGAAACTAC 5756
 Qy 1367 rGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSe 1387
 Db 5757 CAAGACCTCGAAGATGCTGACTATTCTTATGAAATTTATGAGAGACACACGAGACCCC 5816
 Qy 1387 rThrAspGlySerMetAsnSerTySerSerGluGlyAsnLeuIlePheProGlyVal-- 1406
 Db 5817 TGAAGAGGGTGGGTACTCATATGACATAAGTGAAGAGACACACGAGCCCGGAGTGG 5876
 Qy 1407 -----ArgLeuAlaSerAsp-----SerG 1413
 Db 5877 TGGTTACGATGAAAAAGACTGAGAGGTCTAGAGGCTTCTGGATGATCATGCAATGG 5936
 Qy 1413 nPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThr-LeuA 1433
 Db 5937 CTATGATGACTCTGAGGAT-----GGTGGCCACACACTTGG 5972

Qy 1433 laThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGluV 1453
 Db 5973 GGAACCCAGCTACT-----CTTATGAACACCACTGAGAAAATTACCAG 6014
 Qy 1453 alGluIleIleArgAlaArgGlyLeuValLysProGlySerLysThrLeuProAlaP 1473
 Db 6015 TTTCCCTG-----AGTCTGAAGGTATTTCCTATGA 6044
 Qy 1473 roTyTrVal-----LysValTyTrLeuLeuAspAsnGlyValCysIleAlaLysL 1489
 Db 6045 GACATCTACAAACACACACACACCTGATATCTCCACATACTGT----- 6091
 Qy 1489 yslYsThrLysValAlaArgLysThrLeuGluProLeu----- 1501
 Db 6092 -----ACGAGACTGCAGAGAAATCCTAGAACCCCTCAGGCATCCACATATTCCTACGA 6146
 Qy 1502 -----TyTrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuG 1518
 Db 6147 GACTTCAGACCTATGCTACACTGCAG-----AAAAGAAGTCCCTCAGAACCCCGTCA 6200
 Qy 1518 InIleIleValTrpGlyAspTyTrGlyArgMetAsp-----HisLysSerPhe- 1533
 Db 6201 GGATGTGCAATTTATGCTCGTGTCTCTGTGTAATACAAGCACCCCAAGACAGAGCTTTC 6260
 Qy 1534 -----MetGlyValAlaGlnIleLeuLeuAspGluLeu----- 1544
 Db 6261 ACCCTCTTCATTAATCCCAATCTCTTGAGTGGTTTGCCAGTGAAGAACCCACTGAAGA 6320
 Qy 1545 -----ATCTGAAAAGCCCTCACTCAATCAGGGGAGGCCACCGCTCCAGGAGGAAGCAACA 1549
 Db 6321 ATCTGAAAAGCCCTCACTCAATCAGGGGAGGCCACCGCTCCAGGAGGAAGCAACA 6380
 Qy 1549 etValIleGlyTrpPheLysLeuPheProProSerSerLeuValAspPro----- 1565
 Db 6381 GGGCCGACAGTGTGATGAACCCCTCCACCTCAGTCAGCGAGTCAGCCCATCCAGAC 6440
 Qy 1566 -----ThrSerAlaProLeuThrArgArgAlaSerGlns 1577
 Db 6441 CGACTCTGATGTTCCCGGAGAGAGTGAAGAGTGCCTCTCCATCAGCGCGATGCCAATAT 6500
 Qy 1577 exSerLeuGluSerSerThrGlyProSer 1586
 Db 6501 CGACTCTGAAGACAGAGTCCGAAACCATCC 6529

RESULT 10

US-10-252-157-218
 ; Sequence 218, Application US/10252157
 ; Publication No. US20030190640A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paris, Mary
 ; APPLICANT: Pearson, Cecelia I.
 ; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
 ; FILE REFERENCE: PA-0027-1 US
 ; CURRENT APPLICATION NUMBER: US/10/252,157
 ; CURRENT FILING DATE: 2002-10-01
 ; PRIOR APPLICATION NUMBER: 60/295,048
 ; PRIOR FILING DATE: 2001-05-31
 ; NUMBER OF SEQ ID NOS: 501
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 218
 ; LENGTH: 11950
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20030190640A1 475473.1
 ; NAME/KEY: unsure
 ; LOCATION: 9422, 10229, 10231, 10233-10352
 ; OTHER INFORMATION: a, t, c, g, or other
 US-10-252-157-218

Alignment Scores:

Fragment Scores:	
Pred. No.:	5.21e-12
Score:	295.00
Length:	11950
Matches:	366
Percent Similarity:	31.37%
Conservative:	221
Best Local Similarity:	19.56%
Mismatches:	691
Query Match:	595
Indels:	3.56%
DB:	12
Gaps:	78

US-09-617-099B-1 (1-1590) x US-10-252-157-218 (1-11950)

Qy	1	MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro	20
Db	2098	ATGTCAAAACCCAAAGCTGCCAAGGAGAGACG-CTGAAAAAGGAAAAA	2156
Qy	21	ProProGlnProGluMetProAsp-----LeuSerHisLeuThrGluGluArgLys	38
Db	2157	GAAGACAAGAAAGGAGGAAGAAGAAAGCCAAAGAAAGAGTGGCTAAAAAGGAGGACAA	2216
Qy	39	IleIleLeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeu	58
Db	2217	ACACCTATCAAGAGGAGGAGAAACCAACCAAAAGGAGAGGTGAAAGAAAGACGACAAAA	2276
Qy	59	LysIleLysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThr	78
Db	2277	GAGATCAAGAAAGAGAGAAAGAAACCCCAAGAA-----	2312
Qy	79	GluLeuValAsnAenValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGln	98
Db	2313	GAGTTTAAGAAAGAAACACCGCCAAAGAGAGTCAAGAGGAGAGTAAAGAGGAGAGAGAG	2372
Qy	99	ThrLysLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGlu	118
Db	2373	AAGGAAGTGAAGAAAGGAGAAAGGAACCCAAAGAAAGAAATTAAGAGCTCCCTTAAGAC	2432
Qy	119	SerGlnGlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThr	138
Db	2433	GCAAGAAATCATCTACTCTCTCTGAAGCAAA-AAAAACCAGC--TGCTTTAAAAACC	2488
Qy	139	LysPheAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCys----	156
Db	2489	AAAAGTACC-----CAAGAAGGAGAGTCTGTCAAGAAAGATTCTGTGTGCTGC	2536
Qy	157	-----AlaArgCys	159
Db	2537	CGGAAACCCAAAGGAGAGGGGAAATAAAGTCATTAAAGAGAAAGCGCAAGCCGACGA	2596
Qy	160	GlyGlyArgValSerLeuArgSer-----	167
Db	2597	GGCTGTGCTGCACGCTGCGCACTGCGCACTGGAGCCACCACAGCAGCTGTTCATGGCGGACGTGG	2656
Qy	168	AsnLysValMetTrpValCysAsnLeuCysArgLys-GlnGln-----	181
Db	2657	AATAGCAGCCATTGGCCCTGCCAAGAACTCGAAGCTGAGAGGTCCTTTATGTCATCTCC	2716
Qy	182	GluIleLeuThrLysSerGlyAlaIleTrpPheTyrAsnSerGlySerAsnThrLeuGlnI	201
Db	2717	TGAGGATCTAACCAAGGAC-----TTTGAAGAGTTAAAGGC	2752
Qy	201	nProAspGlnLysValProArgGlyLeuArgAsnGluAlaProGlnGluLysLysAl	221
Db	2753	TGAAGAGTGCATGTAAACAAAGGACATCAAGCCCTCAGCTCGAGCTAATCGAAGACGAAGA	2812
Qy	221	AlaLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerVal-----	238
Db	2813	GAAACTGAAGAAACT-----GAGCCAGTCGAAGCCCTACGTCATCCAGAA	2857
Qy	239	----ProAlaValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAs	257
Db	2858	GGAGAGAGAGTCAACAAAGGTCTGCGGAGTCCCTGTAGTGGGAATCTACTACCACCTGA	2917
Qy	257	ngLysGlyValLysHisGlnIleAlaSerAspMet---ProSerAspArgLysArgSer	276
Db	2918	AGGGGAGGGCAATGTGAACAGACACTGAGGAGCTGGAGCCGCTCGAGAAAGCAGGAGT	2977

[illegible]

Db 5525 AGTCAAAAGTCTGGAAGAGAGAGCTCTCTCCAAATCTGATATC----- 5570
 Qy 1216 gSerGlyTrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAs 1326
 Db 5571 -TCTCCACTACCCACAGAGAGTCTCTCTTTATATTCCTACTTTTTCAGATTCT-- 5627
 Qy 1236 pValSerAspValSerAlaValSerArgThrSerSerAlaSerArgPheSerThrSe 1256
 Db 5628 -----ACCTCTGACGTCAGAGAGAAACAGCAACTTGCACAGTTCCTCTTCTCC 5677
 Qy 1256 rYrMetSerValGlnSerGluArgProArgGlyAsnArgLysLysSerValPheThrSe 1276
 Db 5678 ACCAATAGATGACATCCGAGAGCCCTATGGCTTCGCT---GCCTCAGTGTATTTCGA 5734
 Qy 1276 rLysMetGln-----AsnArgGlnMetGlyValSerGlyLysAsnLeuTh 1291
 Db 5735 TACAATGCACACCATCTAGCTTGAATAGAGATTGTCCACACCTGCGCTGGAGAGGA 5794
 Qy 1291 rLysSerThrSerLysSerGlyAspMetCys-----SerLeuGluLysAsnAs 1307
 Db 5795 CAGTGGAGGAGACACCTGGTGACTTAGCTATGCTATCAAAAGCCTGAGAAACAC 5854
 Qy 1307 pGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSe 1327
 Db 5855 CAGTCTCCCATGATGAGAGAGATTATGACTATGACTCTTATGAGAGACACACCGGACCTC 5914
 Qy 1327 rSerLleGlyAlaLysMetValAlaLleValGlyLeuSerArgLysSerArgSerAlaSe 1347
 Db 5915 AGATGGGGTGGCTATTACTATGAGAGATGAGAGAACCAACAAATCTCCAAAGTGACAG 5974
 Qy 1347 rGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerTh 1367
 Db 5975 TGGCTACTCTAT-----GAGACCATTTGGGAAACTAC 6007
 Qy 1367 rGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSe 1387
 Db 6008 CAAGACCCCTGAGATGGTACTATCTCTATGAAATATTATGAGAGACACACCGACCCC 6067
 Qy 1387 rThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnLeuLlePheProGlyVal-- 1406
 Db 6068 TGAAGAGGGTGGTACTATATGATATGATAGTGAAGAGACACACCGCCCCCGAAGTGAG 6127
 Qy 1407 -----ArgLeuAlaSerAsp-----SerG1 1413
 Db 6128 TGGTTACAGCTATCAAAAGACTGAGAGCTTAGAGGCTTCTGATGATCATCAGCATGG 6187
 Qy 1413 nPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThr-LeuA 1433
 Db 6188 CTATGATGACTCTGAGGAT-----GGTGGCCACACACTTGG 6223
 Qy 1433 laThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGluV 1453
 Db 6224 GGACCCAGCTACT-----CTTATGAAACCACTGAGAAATATTACCAG 6265
 Qy 1453 alGluIleIleArgAlaArgGlyLeuValVallysProGlySerLysThrLeuProAlaP 1473
 Db 6266 TTTCCCTG-----AGTCTGAGGTTATTCTCTATGA 6295
 Qy 1473 roTyrVal-----LysValTyrLeuLeuAspAsnGlyValCysIleAlaLysL 1489
 Db 6296 GACATCTCAAAAGACACACGACCCCTGATACTTCCACATCTGTI----- 6342
 Qy 1489 yLysThrLysValAlaArgLysThrLeuGluProLeu----- 1501
 Db 6343 -----ACGAGACTGACAGAAATCACTAGAACCCCTCAGGCATCCACATATTCCTACGA 6397
 Qy 1502 -----TyrGlnGlnLeuSerPheGluGluSerPheGlnGlyArgValValLeuG 1518
 Db 6398 GACTTGACACTATGCTACTGAG-----AAAAGAGTCCCTTCAGAGCCCGTCA 6451
 Qy 1518 lnIleIleValTrpGlyAspTyrGlyArgMetAsp-----HisLysSerPhe- 1533
 Db 6452 GGATGTGATTTATGCTCTGCTCTCTTGTGATATACAGACACCCCAAGACAGAGCTTTC 6511

Qy 1534 -----MetGlyValAlaGlnIleLeuLeuAspGluLeu----- 1544
 Db 6512 ACCCTCTTTTCAATTAATCCCAATCTCTTGTAGTGGTTTGCAGTGAAGAAACCCACTGAAGA 6571
 Qy 1545 -----GluLeuSerAsnM 1549
 Db 6572 ATCTGAAAGCCCTCACTCAATCAGGGGAGCCCAACCGCTCCAGAGAGGAAGCAACA 6631
 Qy 1549 etValIleGlyTrpPheLysLeuPheProSerSerLysValAspPro----- 1565
 Db 6632 GGGCCGACAGTGTGATGAATAACCCCTCCACCTCAGTCAGAGTCAGCCCATCCAGAC 6691
 Qy 1566 -----ThrSerAlaProLeuThrArgArgAlaSerGlnS 1577
 Db 6692 CGACTCTGATGTTCCTCCCGAGACTGAGAGTGCCTTCATCAGCGCGGATGCCAATAT 6751
 Qy 1577 erSerLeuGluSerSerThrGlyProSer 1586
 Db 6752 CGACTCTGAGACGAGTCGAAACCATCC 6780
 RESULT 11
 US-10-132-134-1
 ; Sequence 1, Application US/10132134
 ; Publication No. US20030171562A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Farnet, Chris
 ; APPLICANT: Yang, Xianhu
 ; APPLICANT: Staffa, Alfredo
 ; APPLICANT: Zazopoulos, Emmanuel
 ; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES
 ; FILE REFERENCE: 3012-2US
 ; CURRENT APPLICATION NUMBER: US/10/132,134
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 52101
 ; TYPE: DNA
 ; ORGANISM: Streptomyces platensis subsp. rosaceus
 US-10-132-134-1
 Alignment Scores:
 Pred. No.: 9,6e-11 Length: 52101
 Score: 288.50 Matches: 432
 Percent Similarity: 31.41% Conservative: 201
 Best Local Similarity: 21.44% Mismatches: 697
 Query Match: 3.48% Indels: 688
 DB: 12 Gaps: 94
 US-09-617-099B-1 (1-1590) x US-10-132-134-1 (1-52101)
 Qy 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 Db 7891 CTGGCTGCACCACTG---GTACGTGCGCGCAACGCCCGCGCGGCTCGCTCGGGT 7947
 Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIleIle 40
 Db 7948 CCCCCTGTACTGTCTGCGACCG-----GCGCAGTGGCGCCCTGTGCGC 7989
 Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluGluLysSerValLeuLysIle 60
 Db 7990 CTAGGA-----CGGAGAACGCGCCATGAGAGCAAGACGCGCTCTCGCCGACGA 8040
 Qy 61 LysGluGluHisLysAlaGlnProThrGlnTnPheProPheSerGlyIleThrGluLeu 80
 Db 8041 CCGGGACATCGCGCTCATCGGCTGTC-----CCTGCGGTTCGCCGCTCGCGCAC 8091
 Qy 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnPro-----AsnGluLysGluProGln 98
 Db 8092 GCCCGAGGAGTCTGAGGACCACTGCGCGAGGCGCGCTCGCTCATCAGCGAGTCCCGGA 8151
 Qy 99 ThrLysLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGlu 118

Db 12041 CGGAACACGGCCGCGCGCTCTCGAGGAGGAGCGAAGCTCTAGCCCGTACACAGGAC 12100
QY 1207 -----LeuGluGlnAspTyrHis-SerLysTyrAr 1216
Db 12101 CTGCTGCGGAACCTGTTGCGGAGGTACCTTGCAGGACGGGAGCAATCTTGGCCGAG 12160
QY 1216 gSerGlyTyrAspProHisArgGlyAlaAspThrValSerThrLysSer 1232
Db 12161 GAGAAGCTGGAGACCT-----ACGGTATCGAATCGATCTCCATCGTCGAG 12205
QY 1233 -----SerAspSerAspVa 1237
Db 12206 CTGACACGAGCTGGAGGACACCTTCGGGTGCTGCCAAGACGCTCTTCTTCAGTAC 12265
QY 1237 lSerAspValSerAlaValSerArgThrSerAlaSerArgPheSerThrSerTy 1257
Db 12266 GTGATCTGAGGGGTGGCGGCTACTTGTGCCGAGCACCGGACCGGCTCTCTGAA 12325
QY 1257 rMetSer-ValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerL 1277
Db 12326 CTCTTCCGCCCGAAGACCCGCGCGGAGCACCACCGCCCGGAGC-----ACCGCC 12379
QY 1277 yMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleS 1297
Db 12380 CCGAAGACACCGCGCGCGGAGGCGCGC-----C 12409
QY 1297 exGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyA 1317
Db 12410 CCGAGGGGCTGC-----CGTCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCC 12460
QY 1317 laLeuGlyThrSerGlyLysArgArgSer-----SerIleGlyA 1331
Db 12461 GTCCGCGCGTCCGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12520
QY 1331 laLysMet-ValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeuSer 1350
Db 12521 CGCCACGACATCCGCTCATCGTATGCGGGCGCGGTACCCGCGCGCGCGCGCGCGCGCG 12580
QY 1351 GlnThrGluGlyGlyGlyLysLysLeuArgSerThrValGlnArgSerThrGluThrGly 1370
Db 12581 GAGTCTCG-GGAGCTGCTAG-----CGAGGCGCGCGCGCGCGCGCGCGCGCGCG 12633
QY 1371 LeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThrAspGly 1390
Db 12634 ATCCGCGTGGCGCAGCGGACATCTACTTCGAGGCGGTGAGCTGACGCGGAGACCGT 12693
QY 1391 SerMetAsnSerTyrSerSerGluGlyAsnLeuIlePheProGlyValArgLeuAlaSer 1410
Db 12694 CGTCAAGACCGGACCTTCTGCGGAGCGT-----CGAGGCGTTCGA 12735
QY 1411 AspSerGlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGln 1430
Db 12736 ----TCGCGCTACTTCAACATCTCCAGCGGAGCGCGGAGCTGCTGTCGCGGAGTCCG 12792
QY 1431 ThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysGlyGln 1450
Db 12793 -----GCTGTTCTCCAGCGGG----- 12810
QY 1451 LeuGluValGluIleLeuAlaArgGlyLeuValValPheProGlySerLysThrLeu 1470
Db 12811 -----CGTGA-----GGCCTGGAGGACCGGGCTA 12837
QY 1471 ProAlaProTyrValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLysLys 1490
Db 12838 -----CTCAGTGCAGAGCTGCGCGCGCGCGCTACGACGCGCA 12873
QY 1491 ThrLysValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSerPheGluGlu 1510
Db 12874 CTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12915
QY 1511 SerProGlnGlyArgValLeuGlnIleLeuValTyrGlyAspTyrGlyArgMetAspHis 1530
Db 12916 -----CTACGGCTTCCAGAACATGCTGATCGCGGCGCACCGGACCGGACCG 12957
QY 1531 Lys-SerPheMetGlyValAlaGlnIleLeuLeuAspGluLeuGluSerAsnMetVa 1550
Db 12958 CGGACGCGAGCTCGGTGTG-----ATGCGGACATGCT 12990
QY 1550 lIleGlyTyrPheLysLeuPheProProSerSerLeuValAsp-ProThrSerAlaPro- 1569
Db 12991 GTCTGACCACTACGGCTTACCGGGCGGCTCCGTTCTTCGACACCA-----TGTGCTCCTC 13047
QY 1570 -----LeuThrArgArgAlaSerGlnSerLeuGluSerSerThrGlyP 1585
Db 13048 GCGCTCGGCTGTGTGCACCGAGCGGTGCTATGCTGCGCAGCGGAGTGCCTGATGAC 13107
QY 1585 roSerTyrSerArgSer 1590
Db 13108 CGTCGTCGCGGCGCATCA 13124
RESULT 12
US-10-177-293-448
; Sequence 448, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Zhao, Xumei
; APPLICANT: Ganavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: East Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10177,293
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 448
; LENGTH: 4215
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-448
Alignment Scores:
Pred. No.: 9,22e-12 Length: 4215
Score: 284.50 Matches: 206
Percent Similarity: 35.85% Conservative: 150
Best Local Similarity: 20.75% Mismatches: 345
Query Match: 3.43% Indels: 295
Gaps: 38

US-09-617-099B-1	(1-1590) x US-10-177-293-448	(1-4215)
QY	25 GluMetProAspLeuSerHisLeuThrGluGluArgLysIleValMet	44
Dd	569 AAATGATTGACTTAAGCTTCTGACTGAAGAGAACAAAGAGGCACATCGAAGTTTG	628
QY	45 AspArg-----GlnLysLysGluGluGluLysGluInSerVal---	LeuLys 59
Dd	629 CAGCGGGATGCTGCTCAGAAGAGGCCGGAAGAAGAGAGAGTCAGACATTTCCCTCGAAAAA	688
QY	60 IleLysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGlu	79
Dd	689 ATTAAGATGACAGCA--GCTGAAGATAATGAGTGG---CCAATG-	729
QY	80 LeuValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThr	99
Dd	730 -----GTTTTATGAGCCAGCAAGCAAAAAGGCACAGGNCATAATCATGTGGCG	777
QY	100 LysLeuHisGlnGlnPheGluMetTyrlsGluGlnValLysMetGlyGluGluSer	119
Dd	778 AGATATCATCAGAGCATCTATGAGAAAGAGAGGCCGCCAGATAGCAGCTGAGCAGAGTAA	837
QY	120 GlnGlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrIys	139
Dd	838 AGACAGAGAAAAATGGGGCAAAGCA-----AAG	864
QY	140 PheAlaAspGlyCysGlyHisAsnCysSerTyrcysGlnThrLysPheCysAlaArgCys	159
Dd	865 CTGGGTGAATATGTCAACAAGATGC-----	891
QY	160 GlyGlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLys	179
Dd	892 -----TTTCCTTCCTCCAGAGCTGGCTGGCTGTGAGAGAGGCAGCAAGA	936
QY	180 GlnGlnGluLysLeuThrLysSerGlyAlaTrpPheTyraen-SerGlySerAsnThrLe	199
Dd	937 AGATGACAGCACCC-----AGCAAGCCCGAGTTCACAGTGT	969
QY	199 uGlnGlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLy	219
Dd	970 GGTAATCCAGCTTCAGTGTG-----ATTGATGTCCCAGGAAAA	1011
QY	219 slysAlaLysLeuHisGlnGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValPr	239
Dd	1012 CACAAGAAAACCAGATGTCTCCAGAGAGAGAGAGAAATCGTTTAAATGCTCC	1066
QY	239 oalaValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrLysAsnGly--	258
Dd	1067 -AAGTTGCCAGAGAGTCACTCATCAACAATACTAAAAATGAACAGTCAAAAAATGGAAG	1125
QY	259 -SerGlyValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSe	278
Dd	1126 AACTGTTTATTTTCAGACTTCAAAAGAGATGCAATGTTCAGAGTCAAAAGAAAAGTCAAC	1185
QY	278 rValSerArGspGlnAsnArgTyrgluGlnSerGluGluArg----GluAspTyrSe	297
Dd	1186 TGTCGAGATACTTCAATCCAAAAGTTAGAGAAATCAAAAGCAGACTTTTGCCAGGCCTTC	1245
QY	297 rGlnTyValProSerAspGlyThrMetProArgSerProSerAep-----	312
Dd	1246 AAA-TGGTTCCTCAATCAAGGCTCCAAATCCCCAAGCCAGGAAGATGATCTACAAATCAA	1304
QY	313 -----TyrAl 314	
Dd	1305 CTGATTAAACAAGATGATACCACTCTTTTCTTAGACAAAGGACAGACTCCCTCGAAAG	1364
QY	314 aspArgArgSerGlnArgGluProGlnPheTyrgluGluProGlyHisLeuAsnTyraR	334
Dd	1365 CGAGAGGGCTCCGAGAGGATCTTCAA-----GGCAACTCCAGTTCCA	1409
QY	334 g-AspSerAsnArgArgGlyHisArgHisSerLysGluTyrlleValAspAspGluAsp	354

```
QY 706 heilethrysvallyslyserleuAlaaspThrValGlyHisLeuArgProGlyA 756
Db 2120 -----TCAGATTCAATCAAGCAAAAGTAGATA 2150
QY 726 spGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrA 746
Db 2151 ATCAGCCAGCAAGAA-----TTAGTGGGTAGTGTCTGAAGATGATGAG-----2191
QY 746 snilelleLeuGluSerLysProGlu-----ProGlnValGluLeuValValSerArgP 764
Db 2192 -----AAACCAAGATCAGAGCCAGTTACAAATGAATGCGTACCAAGA- 2233
QY 764 roiledGlyAspIleProAspSerThrHisAla-----GlnLeuGluS 781
Db 2234 --ATTTCACAGGCGCTACACAACTTGATATCCATTCTTCACCTCGCAAACTCAAA 2291
QY 781 erSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSerP 801
Db 2292 GGATGAGCAAGTCTGTTCCAGCATTTCTCCAGATGAGGTGAGTGGCAGTGTGATGATG 2351
QY 801 roMetSerPro-----GlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuS 819
Db 2352 TTATAGTGGAGACTTTGGCAATCTGAA-----GTTAAAGGAAATATTC 2396
QY 819 exileLysLeuTrpPheAspLysValGlyHisGlnLeuLeuValThrIleLeuGlyAlaL 839
Db 2397 AGTTTCAATTGAATATGAGTCACTGAGAGTGTGATGTTTTTGGCCCACTGTA 2456
QY 839 yAspLeuProSerArgGluAspGlyArgProArg---AanProTyrValLysIleTyrP 858
Db 2457 AGGACTTAGCAGCGGATGTAAACAAACAGCGTTTCAGACCCATATGTAAGGCGCTATT 2516
QY 858 heLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuG 878
Db 2517 TGCTACCAACAAAGCAAAATGGGCAAGAGAAACACTCGTAGTGAAGAAACCTTCA 2576
QY 878 luProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluA 898
Db 2577 ATTCCTGTATACGAAATACCTGGGTATAAA---ATTGAACAAACAAATCTTAAGACAC 2633
QY 898 rgMetLeuIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheL 918
Db 2634 AGAAATTGAACCTGTCTCAATTGGCAT-----CGGATACATTTAAGCGCAATGTTCC 2687
QY 918 euGlyLleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHis-----935
Db 2688 TAGGGAGGTGGAACCTTGATTTGGAAACATGGGACTGGGATAACCAACAGATAAACAA 2747
QY 936 -----TrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeuProArg-----951
Db 2748 TGAGATGTTACCTCTGAAGCG- GAAGACAGCAGCAGTGGCCCTTGAAGCAAAACAGA 2806
QY 952 -----ProSerProTyrLeuProArg 958
Db 2807 GGTGAAATGAACACTAGCTCTCCAGTATGTCACAGA 2841
```

RESULT 13

```
US-09-814-353-19153
; Sequence 19153, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
```

```
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19153
; LENGTH: 11295
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11276..11277, 11278, 11279, 11280, 11281, 11282, 11283, 11284,
; LOCATION: 11285, 11286, 11287, 11288, 11289, 11290, 11291, 11292,
; LOCATION: 11293, 11294, 11295
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19153

Alignment Scores:
Pred. No.: 6,65e-11 Length: 11295
Score: 280.00 Matches: 374
Percent Similarity: 32.84% Conservative: 229
Best Local Similarity: 20.37% Mismatches: 671
Query Match: 3,38% Indels: 564
DB: 12 Gaps: 85

US-09-617-099B-1 (1-1590) x US-09-814-353-19153 (1-11295)
QY 31 HisLeuThrGluGluGluArgLysIleIleLeuAlaValMetAspArgGlnLysLysGlu 50
Db 364 CATGGAATCTCAGAGAGAGGTATGTCG-----GTAATCTTCGGCAGCAAGAGGAC 417
QY 51 GluGluLysGluGlnSerValLeuLysLysLys-----GluGluHisLysAlaGlnPro 68
Db 418 CAAGAAAAAGCCCTTGACTGCATCAAAAGGAAAACTTTCTTGGCATGCAGATTGAAGTA 477
QY 69 ThrGlnTrpPheProPheSerGlyIleThrGluLeuValAsnValLeuGlnProGln 88
Db 478 ACAGCATGGATA-----GGTCCAGAAACAGAAAGTGAATAATCAATTTCCGCCCTTG 528
QY 89 GlnLysGlnProAsnGluLysGluProGlnThr-----99
Db 529 GATGAAGAGTAGATGAATTTCCACCCCAAGCAAGAACTCTCTTTATTGGCAACCTT 588
QY 100 -----LysLeuHisGlnGlnPhe-----GluMetTyr 108
Db 589 GAAAAAACCACTACTTACCATGACCTTCGCAACATCTTCCAGCGCTTTGGAGAAATTGTG 648
QY 109 LysGluGlnValLysLysMetGlyGluGlnSerGlnGlnGlnGlnGlnGlnGlnGlnGln 128
Db 649 GATATTGACATTAAGAAAGTAAATCGAGTCTCTCAGTATGCGTTTCTCGCAATCTGTCAT 708
QY 129 AlaProThrCysGlyIleCysHisLysThrLysPheAlaAspGly-----CysGlyHis 146
Db 709 -----ATTGCTAGCGTTTGTAAAGCTATTAAAGATGGATGGGGAATATCTTGGAAAT 762
QY 147 AsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGlyGlyArgValSerLeuArg 166
Db 763 AAT-----CGCTCAAGCTGGGT 780
QY 167 SerAsnLysValMetTrpValCysAsnLeuCysArgLysGlnGlnGlnGlnGlnGlnGln 186
Db 781 TTTCGAAAGAGCATG-----CCTACA 801
QY 187 SerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGlnGln-----201
Db 802 AACTCGGTGGCTAGATGGGCTTTCTTCGAATGTGTGATGATGATTTTAAACACGACAT 861
```

Qy	518	LeuArgAsnAepSerLeuSerAspGlnSerGluSerValArgProProProArg	537
Db	1678	ATACGGAAGCAAAAGTGGAAAAGGACAAAACTGAC	1713
Qy	538	ProHisLysSerLysGlyGlyLysMetArgGlnValSerLeuSerSerSer	555
Db	1714	-----AAGCAGAAACGCAAGAAAGAGTTCACTCCCTAGTTCTCAGTCTTCAGAAACG	1767
Qy	556	-----GluGluGluLeuAlaSerThrProGluTyr-----ThrSerCysAspAspValGlu	572
Db	1768	GACCAAGAAATGAGCGAGAGCAAGCCCTGAAAGCCGAGGAGTTGTAAAT-----AAA	1821
Qy	573	LeuGluSerGluSerValSerGluLysGlyAspSerGlnLysGly	587
Db	1822	CTGAGCAGAGAAAGCTGACAAAGAGGGAATAGCGAAAAACCGCTGGAACTCATGCCT	1881
Qy	588	-----LysArgLysThrSerGluGlnGlyValLeuSerAspSerAsnThrArg	603
Db	1882	TGCGTGTTTGACTCGAGTGAAGAAAGAGGAAAGGTCATTGAC-----CACACTCCCT	1938
Qy	604	SerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGluAspLeuGlu	623
Db	1939	GTGGAAAGTGAAGCCAGCTT-----	1962
Qy	624	TrpSerGluProGlnLeuLysAspSerGlyValAspThrCysSerSerThrThrLeuAsn	643
Db	1963	-----GATAAT-----GACACTGTCAAATCTCTCGCCCTGGAC	1995
Qy	644	GluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAsp	663
Db	1996	CAGAAACTTCAGGTCTCTCAGACGGAGCCCTGC AAAATCTGACTTGTCTAAA-----	2046
Qy	664	ArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAsp	683
Db	2047	-----CTGGAAATCAGTTAGTAATGAA-----GTACCAAGGAA	2079
Qy	684	SerGlyAlaMetLeuGlyLeuLysValValGlyLysMetThrGluSerGlyArgLeu	703
Db	2080	AAGGGGCTTTCAAGCCACTGTTGAAGTGGTG-----GAGAAGGAAGCAGGCTT	2127
Qy	704	CysAlaPheIleThrLysValLysGlySerLeuAlaAspThrValGlyHisLeuArg	723
Db	2128	-----AAAGCCGGAAG-----CACCTCAAG	2148
Qy	724	ProGlyAsp-----GluValLeuGluLutPAsnGly	733
Db	2149	CCTGAGCAGCTGCAGATGGGGTAAGTGTGTGTGATCTGGAGAAGCTGGAGCCAGAAA	2208
Qy	734	ArgLeuLeuGlnGlyAlaThrPheGlu-----	742
Db	2209	AGGCGCTTTCGAGATCCCAATTTAAAGACAGAAAAGCAAAACACAGAGGTCAAAGAAAGC	2268
Qy	743	-----GluValTyrAsnIleIleLeuGluSerLysProGluProGlnVal	757
Db	2269	AGTCCAGATGAGGAGTGTCCCGTGTCTTCAAAAAGCAGCGCTGACGTCTCTTAGA	2328
Qy	758	GluLeuValValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAla	777
Db	2329	GAGTCAATCTGTGAGGAA-----GGAGAGGCTCAAGAAAGGCT-----GTG	2373
Qy	778	GlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSer	797
Db	2374	AGGAAAGAAATCTTAAAGAGAAATCTAAAAAAATCAAACTGGACAGA-----	2421
Qy	798	ValThrSerProMetSer-ProGlyMetLeuArgAspValProGlnPheLeuSerGly	816
Db	2422	CTTAATACTGTGCCAGCCCCAAAGACTGTCCAGAGCTTGCCAGTATTTCTGTGTGGTCT	2481
Qy	817	-----GlnLeuSerIleLysLeuThr	823
Db	2482	GGCTCAAGGCCAGCTCAGACCTACAGCAGACTGGGAGAACCCAGCAGTGAATCTGTG	2541
Qy	823	pPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyValAlaLysAspLeuProSe	843


```
QY 1431 xLeuAlaThrProAlaMetGlyAspIleGlnValGly-MetMetAspLysGlyGlnL 1451
Db 4527 ACTTTCAAAACCAAGCTCAGAAGTCTGAGGAAGCCATGAGCCAAAGCCGAAAGCCAGA 4586
QY 1451 euGluValGluIleIle-----ArgAlaArgGlyLeuValValLysProG 1466
Db 4587 CGCCACTGCAGATGCTGAGCCTGATGCAACACGAGAGCCGAGAGCTGCTCTGAGTCTCA 4646
QY 1466 lysSerLysThrLeuProAlaProTyrValLysValTyrLeuLeuAspAsnGlyValCysI 1486
Db 4647 GCCCCAGCTTCTGAAGATTAGAGGTTGATCTCCAGTTGCTGCAAGG----- 4696
QY 1486 leaLalysLysLysThrLysValAlaargLysThrLeuGluProLeuTyrGlnGlnLeu- 1505
Db 4697 --ATAAAAGCCAAACAAAGCAAGC-----GTTCAAGAGACCCCTGTTGAGGAGCTGC 4748
QY 1506 ----LeuSerPheGluSerProGlnGlyArgValLeuGlnIleIleValTyrpLysA 1524
Db 4749 AGTCAGTATCGTGAGAGCCCGCTCACAGGAGAGTGAG-AGGATA-----G 4795
QY 1524 spTyrGlyArgMetAspHis---LysSerPheMetGlyValAlaGlnIleLeuLeuAsp- 1542
Db 4796 ACCGGGAAAAAAGCTCAAGCGGTCCAATTCTCTCGGGGAGAGACACAGAGCTTTTGGAA 4855
QY 1543 --GluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPheProSerSerL 1562
Db 4856 TGAAGATGGAGGCGCAGAGAGATTACA-----AGGACTGCTTCTAAAGAACTCTG 4903
QY 1562 euValAsp-----ProThrSerAlaProLeu-----ThrArgArgAlaSerG 1576
Db 4904 CTGCAGACCTTGACATCCCGAACCAAGTTTGCTCTCAGCCGNACAGGCGCGGATG 4963
QY 1576 lnSerSerLeuGluSerSerThrGlyProSerTyrSerArgSer 1590
Db 4964 TAAGGAGCGTCTATGCAACCATGGGTGACCATGAAACCGCTCT 5007
```

RESULT 14

```
US-10-177-293-422
; Sequence 422, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzstai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
```

```
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 422
; LENGTH: 12227
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-422

Alignment Scores:
Pred. No.: 8,72e-11 Length: 12227
Score: 279.00 Matches: 374
Percent Similarity: 32.82% Conservative: 228
Best Local Similarity: 20.39% Mismatches: 673
Query Match: 3.37% Indels: 560
DB: 14 Gaps: 86

US-09-617-099b-1 (1-1590) x US-10-177-293-422 (1-12227)

QY 31 HisLeuThrGluGluArgLysIleIleLeuAlaValMetAspArgLysLysGlu 50
Db 1315 CATGGAACCTTCAGAGAGAGGTATGCTCG-----GTATTCTTCGCGACGACAGGAC 1368
QY 51 GluGluLysGluGlnSerValLeuLysIleLys-----GluGluHisLysAlaGlnPro 68
Db 1369 CAAGAAAGACCTTCATCATCAAAAGGAAACCTTTCTTTGGCATGCGAGATTGAAGTA 1428
QY 69 ThrGlnTyrPheProPheSerGlyIleThrGluLeuValAsnValLeuGlnProGln 88
Db 1429 ACAGCATGATA-----GGTCCAGAAACAGAAAGTGAAATGAATTCGCCCTTG 1479
QY 89 GlnLysGlnProAsnGluLysGluProGlnThr----- 99
Db 1480 GATGAAAGGATAGATGATTAATTCACCCCAAGCAAGAACTCTCTTTATTGGCAACCTT 1539
QY 100 -----LysLeuHisGlnGlnPhe---GluMetTyr 108
Db 1540 GAAAAAACCACTACTTACCATGACCTTCGCAACATCTTCGACGCTTTGGAGAAATGTG 1599
QY 109 LysGluGlnValLysLysMetGlyGluGluSerGlnGlnGlnGlnGlnGlnGlyAsp 128
Db 1600 GATATTGACATTAAAGAAAGTAAATGAGTTCCTCAGTATCGCTTCTGCAATCTGTGAT 1659
QY 129 AlaProThrCysGlyIleCysHisLysThrLysPheAlaAspGly-----CysGlyHis 146
Db 1660 -----ATTGCTAGCGTTTGTAAAGCTATTAAAGAGATGGATGGGGAATATCTTGGAAAT 1713
QY 147 AsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGlyGlyArgValSerLeuArg 166
Db 1714 AAT----- 1731
QY 167 SerAsnLysValMetTyrValCysAsnLeuCysArgLysGlnGlnGlnLeuThrLys 186
Db 1732 TTTGGAAGAGCATG-----CCTACA 1752
QY 187 SerGlyValatPheTyrAsnSerGlySerAsnThrLeuGlnGln----- 201
Db 1753 AACTGCGTGGCTAGATGGCTTCTTCGAAATGTGTGATCAGTATTTAAACAGCAT 1812
QY 202 -----ProAspGlnLysValPro-----ArgGlyLeuArgAsn 212
Db 1813 TTCTGCGGATATGGGCTGTGGTAAAGGTGTGTTTGACCGCTTAAAGGCGCATG----- 1866
QY 213 GluGluAlaProGlnGluLysLysAlaLysLeuHisGlnGlnProGlnPheGlnGlyAla 232
Db 1867 -----GCCCTGGTTCTCTACATGAATGAATGATATGACAAAGCA 1905
QY 233 ProGlyAspLeuSerValProAlaValGlu-----LysGlyArgAlaHisGly----- 248
Db 1906 -----GCTGTAAAGAGACCAAGGAGGAAATCGGTGGGAAT 1944
```


QY 249 -----LeuThrArgGlnAspThrIle----- 255
 Db 1945 AAAATTAAGTGATTTTGCAAATCGGAAAGTCAGCTGCCTTTTATCACTCATCGTAGG 2004
 QY 256 LysAsnGlySerGlyVallys-----HisGlnIleAlaSerAspMetProSerAspArg 273
 Db 2005 AAATCTGGTCAACACATCAGAGACTTTTATGAATGTAGCCGAAAGAGAGAGAACGA 2064
 QY 274 Lys-----ArgSerProSer 278
 Db 2065 AGGCGATCTACGACTATTAACCAAGATCGTACATATTATGAGAGTGTTCGAATCCGAGC 2124
 QY 279 ValSerArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGln 298
 Db 2125 ACTTATCTGAGGATCCAGGCGGACTATCCAGCTCGAGGAGAGAGTTTATTCAGAA 2184
 QY 299 TyrValProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSer 318
 Db 2185 TGGGAAACTTACCAAGCA-----GACTACTATGAATCAGCA--- 2220
 QY 319 GlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArg 338
 Db 2221 -----TACTACGATGATCT-----CGGGAATACAGGGAATACAGG--- 2256
 QY 339 ArgGlyHisArgHisSerLysGluTyrIleValAspGluAspValGluSerArgAsp 358
 Db 2257 -----AATGAT 2262
 QY 359 GluTyrGluArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsn 378
 Db 2263 CTTTATGAACAAGATATTAGGAATATAGTTACAGGCAAGGGAACGA----- 2310
 QY 379 LeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGlu 398
 Db 2311 -----GAAAGAGAAAGTGAAGATTGAG 2334
 QY 399 ValSerArgAlaArg-----HisGluArgArgHisSerAspValSerLeuAlaAsnAlaGlu 417
 Db 2335 TCTGACCGGACAGACACATCAGAGGAG-----CCGATTGAA 2373
 QY 418 LeuGluAspSerArgIleSerLeuLeuMetAspArgProSerArgGlnArgSerVal 437
 Db 2374 CGAAGTCAAAAGTCTCTCACTTGGCAGCTCCACAGAGCTCTGGAGCGTCTCCCTCTCAG 2433
 QY 438 SerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGlu 457
 Db 2434 CGAGAGAGTTCCCGAGTGATCTCGAGAGAGGCTTTACAGC----- 2475
 QY 458 AlaGlnGlyGlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArg 477
 Db 2476 -----CGATCTCTACACCGGAGTGGAGCTGTAGTCA 2508
 QY 478 ArgSerProIleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGln 497
 Db 2509 CTCTCCCTCCCAAGATATACAGAACTGGACAAGTCTCGTTTGGACGCTATACAAAAAT 2568
 QY 498 HisHisLeuAspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMet 517
 Db 2569 GAAAAGACAGATAAAGACAGAACTTTTGTATCCGGAGAGAGTGGAGAGAGAGAGAGCTTA 2628
 QY 518 LeuArgAsnAspSerLeuSerSerGlnSerGluSerValArgProProProProArg 537
 Db 2629 ATACGGAAGGAAAGTGGAAAGACCAAACTGAC----- 2664
 QY 538 ProHisLysSerLysLysGlyLysMetArgGlnValSerLeuSerSerSer 555
 Db 2665 -----AAGCAGAAACGCAAGAAAGGTTCATCTCCCTAGTTCTCTGCTTCAGAAAG 2718
 QY 556 -----GluGluGluLeuAlaSerThrProGluTyr-----ThrSerCysAspValGlu 572
 Db 2719 GACCAGAAATAGCCGAGAGCAAGCCCTGAAAGCCCGAGGATTGTAAT-----AAA 2772

QY 573 LeuGluSerGluSerValSerGluLysGlyAspSerGlnLysGly----- 587
 Db 2773 CTGACGACAGAGAAAGCTCACAAGAGGGAATAGCGAAAAACCGCTGGAACTCATGCGCT 2832
 QY 588 -----LysArgLysThrSerGluGlnGlyValLeuSerAspSerAsnThrArg 603
 Db 2833 TCGCTGGTTTTCAGTTCGAGTGAAGAGAAAGAGGGAAGGTTCATTGAC---CACACTCT 2889
 QY 604 SerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGluAspLeuGlu 623
 Db 2890 GTGGAAGATTGAAAGCCCAAGCTT----- 2913
 QY 624 TrpSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThrThrLeuAsn 643
 Db 2914 -----GATTAAT-----GACACTGTCAAAATCTTCTCCCTGGAC 2946
 QY 644 GluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAsp 663
 Db 2947 CAGAAACTTCAGGTCTCTCAGACGCGAGCTCGAAAATCTGACTTCTCTAAA 2997
 QY 664 ArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAsp 683
 Db 2998 -----CTGGAATCAGTTAGATGAAA-----GTACCAAGGAA 3030
 QY 684 SerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSerGlyArgLeu 703
 Db 3031 AAGGGCTTTCAGGCATGTTGAAGTGTG-----GAGAAGGAAGCGAGGCTT 3078
 QY 704 CysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGlyHisLeuArg 723
 Db 3079 -----AAGCCAGGAAG-----CACTTCAG 3099
 QY 724 ProGlyAsp-----GluValLeuGluTrpAsnGly 733
 Db 3100 CCGTACGACCTGCGAGTGGGTAAAGTCTGTGATCTGGAGAGCTGGAGCCAGGAAA 3159
 QY 734 ArgLeuLeuGlnGlyAlaThrPheGlu----- 742
 Db 3160 AGGCGCTTTCAGATCCCAATTTAAAGCAGAAAGCAAAACAGAGGTCAAGAAAAGC 3219
 QY 743 -----GluValTyrAsnIleIleLeuGluSerLysProGluProGlnVal----- 757
 Db 3220 AGTCAGAGATGGAGGATGCTCGCGTGTTCATAAAAGAGCGCTGAGTGTCTCTAGA 3279
 QY 758 GluLeuValValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAla 777
 Db 3280 GAGGTCAATTCGTGAGGGA---GGAGAGCTGAAAGAAAGCCT-----GTG 3324
 QY 778 GlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArg----- 793
 Db 3325 AGGAAGAAATCTTAAAGAGAAATCTAAATAAATCAAACTGGACAGACTTAATACTGTT 3384
 QY 794 -----ProSerIleSerValThrSerProMetSerPro 804
 Db 3385 GCCAGCCCAAGAGCTGTGAGGAGCTTCCAGTATTTCTGTGGTCTGGCTCAAGGCC 3444
 QY 805 GlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIleLysLeuTrpPhe 824
 Db 3445 AGC-----TCAGACCTACAGCAAGACTGGGA---GAATAGCAGGT 3483
 QY 825 AspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArg 844
 Db 3484 GAATCTGTGGAATAATCAA-----GAAGTCCAATCAAAAAAGCCCAATTCCTCAAAA 3534
 QY 845 GluAspGlyArgProArgAsnProTyrVallys---IleTyrPheLeu----- 859
 Db 3535 -----CCACAGCTCAACAGCTGCAGGTATAGATGATCAAGGA 3573
 QY 860 ProAspArgSerAspLysAsnLysArgThrLysThrValLysLysThrLeuGluPro 879
 Db 3574 CCAGAGAGAGAGACGTTAGGAAAAAATATTGTCAGTCTCTGTCAGTGAACACCTGAACGT 3633
 QY 880 LysTrpAsnGlnThrPheIleTyrSerPro----- 889


```
Db 1852 GTCTCTCTCAAGCGCGCGGAGCGCGGCGACCCCGCGC---CGGAGCCCTCTCGCTGACC 1908
|||
Qy 351 spGluAspValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnA 371
|||
Db 1909 CTGAAGGC-----GCCGACGCGGTGTCACCGCGCGCGTCTGAGCGCGATGTG 1962
|||
Qy 371 laArgTyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluG 391
|||
Db 1963 GACGGGGCGCGCTCAACCGCGTCTCTTCGC-----CGCCACAGTCTGCG 2007
|||
Qy 391 luGlnMetArgIleHisAlaGluValSerArgAlaArgHisGluArgHisSerAspV 411
|||
Db 2008 AACCCCAAGTCGCGGACCTCGGTTCGCCCGCACCGCCCGTACAGGTCTCCGACAC- 2066
|||
Qy 411 alSerLeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArg 431
|||
Db 2066 ----- 2066
|||
Qy 431 roSerArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrS 451
|||
Db 2067 --CATCGCAAGACAGTACCGCGCGGTACCGCCCTCGCGCTACAAGCGGTCTGTCATGCAT 2124
|||
Qy 451 erMetGluArgThrArgGluAlaGlnGlnSerSerTyr-----ProGlnArgT 468
|||
Db 2125 CGCGCGCGCGGACCGCGGCGCACCGCGCGTCCACCGGACCGCCCGCGCGCG 2184
|||
Qy 468 hrSerAsnHisSer-----ProProThrProArgArg- 478
|||
Db 2185 GAGACTCCACCGCTCTCGCGGACTACGCGATCGTCCACTACAAGCGCGCGGACGCG 2244
|||
Qy 479 -----SerProIleProLeuAspArgProAspMetArgArgAlaAspSerL 494
|||
Db 2245 GACTACACCGGCTGCGCGCTACGCGTGGCGGATCTGGCGGACCGCGGAGTCTGACGAC 2304
|||
Qy 494 euArgLysGlnHisIleuAspProSerSerAlaValArgLysThrLysArgGluLysM 514
|||
Db 2305 TGGCCCGCGGCGAGCTTCTGTCGCGGAGCGGTACGCGGCTTCTGCTCTAGTCAAG 2364
|||
Qy 514 etGluThr-----MetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValA 532
|||
Db 2365 CTGAAGCGCGCGCTCACCGTGACT-TCTCGTCTATCAACGAGGCGGACAGGA 2423
|||
Qy 532 rgProProPro-----ProArgProHisLysSerLysLysGlyL 546
|||
Db 2424 CGTCTCGCGCGACCGACGATCGACGTCAACGAGCGCGGAGGTCTGGTCTGAGCAGG 2483
|||
Qy 546 ysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAlaSerThrProGluTyrT 566
|||
Db 2484 CAAGGAGACCGTACGGACGAGCGCGGACTACCCCGCGGACGACAGACCAAGCGCGT 2543
|||
Qy 566 hrSerCysAspAspValGluLeuGluSerGluSerValSerGluLysGlyAspSerGlnL 586
|||
Db 2544 CATCC-----ACTACACCGCGCGGACGCGGACCTACCGCTGGGAGTCTGACGT 2594
|||
Qy 586 ysGlyLys-ArgLysThrSerGluGlnGlyValLeuSerAspSerAsnThrArgSerGlu 605
|||
Db 2595 CTGACCGGCGCGCGACGCC-----CACGGACTGTGTCCAAGCCCTTGGAGCGCGTGA 2651
|||
Qy 606 ArgGlnLysLysArgMetTyrTyrGlyGly-----His-SerLeu----- 618
|||
Db 2652 CGACGCTTATGGCGCGTCTTCGAGGTCCCGTCAACGACCGCGCCACGACCTCAGCTA 2711
|||
Qy 619 -----GluGluAspLeuGluThrSerGluProGlnIleLysAspSe 632
|||
Db 2712 CATCATCCACAGGCGGACGAGAGGACCTC-----TCCGCGGACCGGTCTGCTGCACT 2765
|||
Qy 632 rGlyValAspThrCysSerSerThrThrLeuAsnGlnGluHisSerHis----- 648
|||
Db 2766 CACGCGCGGACGCGCATGAGGTGTGCTGTGAACGCGGACGAGGACCACTCTGCTGCCGA 2825
|||
Qy 649 -----SerAspLys 651
|||
```

```
2826 GCCCGCGGCGAGCGCGCGCTCTCGACCTGACCACTCCAGCGCGTCTGGAATCGACCG 2885
651 sHisProValThrTriPdnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLe 671
|||
2886 GAACACCGTCTCGCTTGAACCGCTCC-----GACGCGCGCGCTCCACCGAGCTGCT 2936
|||
671 uAsnLysArg-----LeuLysAspGlySerValProArgAspSerGlyAl 686
|||
2937 GTCTCTCGCGAGCGGTCTGATCCGCTCAAGGACGCTCTGCTACCGACGACGACGCG 2996
|||
686 aMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSerGlyArgLeuCysAlaph 706
|||
2997 GTGCTG----- 3003
|||
706 eIleThrLysValLysGlySerLeuAlaAsp-----ThrValGlyHisLe 722
|||
3004 -----CGGCTTTCGAAGACACCGCTCACCGGCGCGCAGAGCGCGCTTCGCGACCT 3056
|||
722 u-----ArgProGlyAspGluValLeuGlu----- 730
|||
3057 GAAGTCGTACACCGCTCGGTCTCGACCCACCGACCGCGCGGTGCGCGAGGTCT 3116
|||
731 -----TrpAsnGlyArgLeuLeuGlnGlyAlaLath 740
|||
3117 CGCGCGCGCGTCTGCTCGCAACGAGCGCGGACGAGCGGTCTG-----GCCGCGAC 3173
|||
740 rPheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVa 760
|||
3174 GGGCGTCTCAGTCTCGCGCGTCTGCGACGATCTGTACGACGCGACGAGCGGCGCTC-- 3231
|||
760 lValSerArgProIle-----GlyAspIleProArgIle-----ProAspSe 774
|||
3232 -----GGCCCAACGTTCCGCGCGCGGCGCACCTCGCGGTGTGCGCGCGCCACCG 3284
|||
774 rThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArgPr 794
|||
3285 GCAGTCTGCTCTGCGTCTGAGTCTGACGCGCGCGCGCGATGAGCGGACACGCGCAC 3344
|||
794 oSerile---SerValThrSerProMetSer----- 803
|||
3345 CGGCGTCTGCTCTGCTCTACCGCGCGCGCTCTCTGAGGCGCAAGCGGTACCGTACGTCGT 3404
|||
804 -----ProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSe 819
|||
3405 GAAGGTGTGCGCGCGCGCGCGTCTCGGAGGTCTGTCACCAACAGGTACCGACCGCTACTC 3464
|||
819 rIleLysLeuTyrPheAspLysValGlyHisGlnLeuIleVal----- 833
|||
3465 GTGCGCGCTCACACCGACTCC-----GAGCGCGCGCTCTGCTGACCTGACGACGAGTGC 3521
|||
834 -----ThrIleLeuGlyAlaLysAspLeuProSerArgGluAs 846
|||
3522 CCTCGCGCGCGCGGTCTGCTGCTGAGAGAGCGCGCGCGCTGCGCTGCGG---GA 3578
|||
846 pGlyArgProArgAsnProTyrValLysIleTyrPheLeuProAspArg----- 862
|||
3579 CGCGGAGATCCAGGAGTGCACATCCGCGACTTCTCGGTGCGGACGAGCGCGTCCGCGC 3638
|||
863 -----SerAspLysAsnLysArgArgThrLysThrVa 873
|||
3639 GAAGGACCGCGGCGACCTACCTCGCTTCCGACGAGAACAGCGACGCGCTCCCGGACCT 3698
|||
873 lLysLysThrLeuGluProLysTrpAsn-----GlnThrPheIleTyrSe 888
|||
3699 CGCGCACTGGCGAGTCTGGGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3758
|||
888 rProValHisArgArgGluPheArgGluArgMetLeuGluIleThrTrpAspGlnAl 908
|||
3759 CACGATCGCGAGAGAGGTCCGCGGACGAGGCGCGCGCTGCGGACTCTGCTGCTGCTGCT 3818
|||
908 aArgValArgGluGluSerGluPheLeuGlyGluIleLeuIleGluLeuGluThrAl 928
|||
3819 GCGCGACTCCGAGAGCAGCAGGAGTCTCTACCGCGGTCTGCGCGCGGAG----- 3867
|||
```

QY	928	aLeuLeuAspAspGluProHisTrpTyrLysLeuGlnThrHisAspValSerSerLeuPr	948
Db	3868	 : : : : : : -GAGCGTACAACTGG-----GGTACGACCCCTTACCACATACAC	3905
QY	948	oLeuProProSerProTyrLeuProArgArgGlnLeuHisGlyGluSerProThrAr	968
Db	3906	: : : : : : : GGTCCCGGAGGCTCG-----TACGCCACCGACCGACCGACCGCG	3947
QY	968	gArgLeuGlnArgSerLysArgLleSerAspSerGluValSerAspTyrAspCysGluAs	988
Db	3948	: : : : : : : : : GCGCACGGTCGAGTTCGCGAGGATGGTCAAGTCGCTG-----AACCAGGA	3992
QY	988	pglyValGlyValValSerAsp-----TyrArgHisAsnGlyArgAspLeuGlnSerse	1006
Db	3993	: : : : : : : CGGCGCTCGGGTCGTATGACGTGCTACAACACACCGCGCGCGCCGCGACGCCGG	4052
QY	1006	rThrLeuSerValProGlnValMet-----SerSerAsnHisCysSerProSerG	1024
Db	4053	: : : : : : : : : CACC---TCGGTGTCTCAGCCGGATCGTCCCGGCTACTACAGCGGCTGTCTGCCACGG	4109
QY	1024	lySerProHisArgValAspValle-----GlyArgThrArg-----	1036
Db	4110	GTCCGTGGCGACGAGCACCTGTTGCGGGAACCGCCACCGACGACGCGATGATGGCAA	4169
QY	1037	-----SerTrpSerProSerAlaProProGlnArgAsnValGluGlnGlyHisArgG	1055
Db	4170	: : : : : : : GTCGTGCTGAGACTCCCTCTGCTCAGCTGGCCA-----AGGA	4205
QY	1055	lyThrArgAlaThrGlyHisTyrAsnThrIleSerArgMetAspArgHisArgValMetA	1075
Db	4206	: : : : : : : GTACAAGTCGACGGCTTCGCTTCGACC-----	4234
QY	1075	spAspHisTyrSerSerAspArgAsp-----ArgAspCysGluAlaAlaAspArgG	1092
Db	4235	: : : : : : : -TCATGGGGCACACAGCCCAAGGCCAATCTCGCGTCCGGAAGCCCTCGACGCGCT	4292
QY	1092	lnProTyrHisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThrA	1112
Db	4293	: : : : : : : CACCGTCGGAAGACCGCGCTCGACGGGAGAAGA-----TCATCTGTACCGGGA	4343
QY	1112	rgSerArgSerSerGluArgProAspThrAsnLeuMetArgSerMetProSerLeuMetT	1132
Db	4344	: : : : : : : GGGTGAAGACTTCGGCGAGGTTCGCCGACGACCGCGCTTCGTGCAGGCCACCCAGAA	4403
QY	1132	hr-----GlyArgSerAlaProProSerProAlaLeuSerArgSerHisProArgThrG	1150
Db	4404	: : : : : : : CATGGCGGCACCGGCATCGCACTTCTCCG-----ACCGGCCCGCGACGCG	4451
QY	1150	lySerValGlnThrSerProSerSer-----ThrProGlyThrGlyArgArgGlyArgGlnL	1169
Db	4452	: : : : : : : CGTACGGGAGGCGGCCCTTTCAGCCGACCCCGGTGTCAGGGCTTCGGGTTCGGGGCT	4511
QY	1169	euProGlnLeuProProLysGlyThrLeuGluArgSerAlaMetAspIleGluGluArgA	1189
Db	4512	: : : : : : : CTACACGACCCCAACTCTCTCGAGCG-----CGAACGGCACA-CGCGCCGACGACGAAGG	4564
QY	1189	snArgGlnMet-----LysLeuAsnLysT	1197
Db	4565	: : : : : : : CCGCGCTGCTGCATCTACAGGACTCATCAAGGTGGGGCTGAGCGGCAACTCTCGCGAAGT	4624
QY	1197	Yr-----LysGlnValAlaGlySerAspProArgLeuGluGlnA	1210
Db	4625	: : : : : : : ACCGCTTCCGACTCCAGCGGCAGAGAGTTCAGGGCTCC-----GAGGTGCG	4672
QY	1210	spTyrHisSerLysTyrArgSerGlyTrpAspProHisArgGlyAlaAspThrValSerT	1230
Db	4673	: : : : : : : ACTACAACGGCAGC---GGGGCGCGCTAC-----GGCGCGCGCCCGCGCG	4714
QY	1230	hrLysSerSerAspSerAspValSerAspValSerAlaValSerArgThrSerSerAlas	1250
Db	4715	: : : : : : : ACGCCCTCGCTACGCGCAGCGCAGCAGC-----AACGAGT	4750

Qy	1250	erArGpheSerSerThrSerTyrMetSerValGlnSerGluArProArGdiYAsnArgL	1270
Db	4751	CCCTGTACGACGGCTCACCTAC---AAGCTCGCGAAGGGGACCCCGGGCGGACCGGG	4807
Qy	1270	ys-----IleSerValPheThrSerLysMet---GlnAsnArGlnMetGlyV	1285
Db	4808	CCCGGATGCAGTCTCGCCACTCGGACGCGCCCTCGCGACGGACCGCTCCTCCC	4867
Qy	1285	alSerGlyLysAsnLeuThrLysSerThrSerLysSerGlyAspMetCysSerLeuGluL	1305
Db	4868	AGCGGGCTCCGACCTGCTCGCTCCAAAGTCCCTC-----GACCGCACTCTACGACA	4921
Qy	1305	ysAsnAsp-GlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSer-GlyLysI	1324
Db	4922	CGCGCGACTCTGGTTCAACGCCATCCACT-----GGAATCGCAGGACGGCA	4966
Qy	1324	sArq-----ArgSerSerileGlyAlaLysMetValAlaIleValClyLeuSerAr	1341
Db	4967	ACGCTTCGGACCGGGACTCCGATGCGCGCCGACCAACAGTCCAAGTGCCTACGCCA	5026
Qy	1341	gLySerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSe	1361
Db	5027	C-----ACCGCTGTGACCTCCCTGAAGGTGG-----CTGCGACC	5062
Qy	1361	rThrValGlnArg-SerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetTh	1381
Db	5063	A-----GATCGAGGGCACTCGCGCGGTACACGAGCACTGCTGAGGATCCGTACGACC	5116
Qy	1381	rgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnL	1401
Db	5117	AGCGGACTTCTCGCTCTCCACGCGCGCCAGGTCCAGTCG-----AAGC	5161
Qy	1401	eullePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyL	1421
Db	5162	TCACCTTCCGCT-CTCCGGCAAGGACGACGACCCCGG---AGTGATCACCATGAGCTC	5217
Qy	1421	euGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleG	1441
Db	5218	GGCGACCTCGCTCGTCTTCAACGGGAC-----GCCCGACCAGCAGGACGAC	5267
Qy	1441	lnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArg	1459
Db	5268	-----GGTCGC-----CGCGCTCGCGCGG	5286
Qy	1460	--GlyLeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrL	1479
Db	5287	AAGGACTACGCTGCACCCCGGTGCAGCGCGGGCGCGCACCTATTCGTCAAGTCCGCT	5346
Qy	1479	euLeu 1480	
Db	5347	TCCTA 5351	

Search completed: November 16, 2003, 20:21:24
Job time : 18945 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 09:17:21 ; Search time 185 Seconds

(without alignments)

3793.509 Million cell updates/sec

Title: US-09-617-099B-1

Perfect score: 8285

Sequence: 1 MSAPLGRGRPAAPTAASQP.....TTRASQSSLESSTGSPYSRS 1590

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DSV=xlh

-Q=/cgn2_1/USPTO.spool/US09617099/runat_14112003.191547_23707/app_query.fasta.1.1735

-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi

-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09617099 -CGN 1 1 95 @runat_14112003.191547_23707 -NCPUS=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCUTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	360.5	4.4	9511	1	US-08-056-200-93
2	360.5	4.4	9511	2	US-08-800-644-93
3	311	3.8	6755	3	US-08-931-999-4
c 4	292.5	3.5	9880	3	US-08-680-897-1
5	290	3.5	4483	4	US-08-961-527-363
c 6	288.5	3.5	33529	3	US-09-144-085-3
c 7	283.5	3.4	8878	1	US-08-759-444-2
8	270.5	3.3	28958	1	US-08-258-261B-6
9	270.5	3.3	28958	1	US-08-456-837-6
10	270.5	3.3	28958	1	US-08-457-342-6
11	270.5	3.3	28958	1	US-08-457-646A-6
12	270.5	3.3	28958	1	US-08-458-076A-6

13	270.5	3.3	28958	1	US-08-764-233A-4	Sequence 4, Appli
14	270.5	3.3	28958	1	US-08-457-335A-6	Sequence 6, Appli
15	270.5	3.3	28958	1	US-08-729-214-6	Sequence 6, Appli
16	270.5	3.3	28958	3	US-09-028-934-6	Sequence 6, Appli
17	270.5	3.3	49377	1	US-08-764-233A-1	Sequence 1, Appli
c 18	266.5	3.2	80161	3	US-09-036-987A-2	Sequence 2, Appli
c 19	266.5	3.2	80161	3	US-09-370-700-1	Sequence 1, Appli
c 20	266.5	3.2	80161	3	US-09-603-207-1	Sequence 1, Appli
c 21	266.5	3.2	80161	3	US-08-452-655B-1	Sequence 1, Appli
22	265.5	3.2	8532	1	US-08-450-582-1	Sequence 1, Appli
23	265.5	3.2	8532	3	US-09-105-537-1	Sequence 1, Appli
24	264.5	3.2	15872	3	US-09-410-551B-1	Sequence 1, Appli
c 25	264	3.2	17536	2	US-08-804-227C-13	Sequence 13, Appli
c 26	263	3.2	13987	2	US-08-804-227C-7	Sequence 7, Appli
c 27	263	3.2	44377	2	US-08-804-198-1	Sequence 1, Appli
c 28	263	3.2	44377	2	US-07-741-940-1	Sequence 1, Appli
29	261.5	3.2	9606	1	US-08-289-548A-1	Sequence 1, Appli
30	261.5	3.2	9606	1	US-08-452-654-1	Sequence 1, Appli
31	261.5	3.2	9606	2	US-08-370-235A-1	Sequence 1, Appli
32	261.5	3.2	9606	2	US-08-449-731-1	Sequence 1, Appli
33	261.5	3.2	9606	3	US-09-335-409-1	Sequence 1, Appli
c 34	260.5	3.1	68750	4	US-09-567-969-1	Sequence 1, Appli
c 35	260.5	3.1	68750	4	US-09-568-102-1	Sequence 1, Appli
c 36	260.5	3.1	68750	4	US-09-568-486-1	Sequence 1, Appli
c 37	260.5	3.1	68750	4	US-09-568-486-1	Sequence 1, Appli
c 38	260.5	3.1	68750	4	US-09-568-486-1	Sequence 1, Appli
c 39	260.5	3.1	68750	4	US-09-568-486-1	Sequence 1, Appli
c 40	260.5	3.1	68750	4	US-09-567-899-1	Sequence 1, Appli
c 41	259.5	3.1	47981	4	US-09-679-279-1	Sequence 1, Appli
c 42	258.5	3.1	32768	4	US-08-961-527-71	Sequence 71, Appli
43	254	3.1	5200	4	US-08-978-277A-3	Sequence 3, Appli
c 44	252.5	3.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 45	250	3.0	50937	3	US-09-428-517-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-056-200-93

; Sequence 93, Application US/08056200

; Patent No. 5616500

; GENERAL INFORMATION:

; APPLICANT: Steinert, Peter M.

; APPLICANT: Lee, Seung-Chul

; APPLICANT: Kim, In-Gyu

; APPLICANT: Chung, Soo-Il

; APPLICANT: Park, Sang-Chul

; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and

; TITLE OF INVENTION: Methods of Using Same

; NUMBER OF SEQUENCES: 117

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive, Sixteenth Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/056,200

; FILING DATE: 30-APR-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fedrick, Michael F.

; REGISTRATION NUMBER: 36,799

; REFERENCE/DOCKET NUMBER: NIH054.001A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (714) 760-0404

; TELEFAX: (714) 760-9502

6022 CGCAGACACATCCCGACAGATCTGCAGCAGCTGCTGGTGAACAGACAGAGAGAT 6081
1264 -----ArgProArggly-----AsnArglylleSerVa 1273
6082 CGTGACCAAGAGAGAGAGCGCTGGCAGCAGGCGCAACAGGCAATTCACAGAGGAAGACAG 6141
1273 lPheThrSerIysMetGlnAsnArgGlnMetGlyValSerGlyIysAsnLeuThrLysSe 1293
6142 CTGGAGCGAGGAAGCAAAAGGAAGCCAAAGGCGCGACAGGAAGTCCCAAGAGGAAG 6201
1293 rThr-----Se 1295
6202 CAGTTCTCAGAGAGGAAGAGAGAGAGACGCCCTCAGAGACACACAGAAAATTC 6261
1295 rIleSerGlyAspMetCys----- 1301
6262 CGCAGAGGAACAGCTGCTCCAGGAAGGAGGAGAACAGCCGCTGCTCGGCCAAGAGCGT 6321
1301 ----- 1301
6322 GACAGAAAATTCGCGAAGAGAGAACTGCTTCATCAGGAACAAAGGAGAGAAAATTCCTCGAG 6381
1302 -----SerLeuGluLysAsnAspGlySe 1309
6382 GAGGAACAGCGCTCGCGGAGGAACCGGAGAGAAAATTCCTTAAGAGAGAACAGCAGCTG 6441
1309 rGlnSerAspThrAlaValGlyAlaLeuGlyThr-----SerGlyLysLysArgAr 1326
6442 CGCCTCGAGGAGCGGAGCACTGGCTCAGGACCGCGACAGAAAATTCGCGAGGAGGAA 6501
1326 gSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAl 1346
6502 CAGCAGCTGAGCGCGCAAGAGCGTGACAGAAAATTCCTGTAAGAGGAGAACAGCAGTGGC 6561
1346 sSerGln-----LeuSerGlnTh 1352
6562 CGCCAGGACGAGAGAGAAAATTCCTGGAGGAGAACAGCAGCTGGCGCAGGAGGTAC 6621
1352 rGluGlyGlyLysLys----- 1358
6622 AGAAAATTCGCGAAGAGGAACAGCTGCTCCAGGAAGGAAGAACAGCAGCTGCACCGC 6681
1359 -----LeuArgSerThrVa 1363
6682 CAAGAGCGTGACAGAAAATTCCTGGAGGAGGAACACAGCTCGCGCGCAAGAGCGTGAC 6741
1363 lGlnArgSerThrGluThrGlyLeuAlaVal-----GluMetArgAsnTrpMetThrAr 1381
6742 AGAAAATTCGCGAAGAGGAACAGCTGCGCAGTCAAGAACAGAGAGAAAATTCCTCGAGAG 6801
1381 gGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnLe 1401
6802 GAACAGCAGCTGCACCGCCAGCAACGCGCAGAGAAAATTCCTC-CAGGAGGAACAGCAGCT 6860
1401 uIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLe 1421
6861 GCGCGC-CAGAGAGCGCGCGCAACAGCGCGTCA----- 6893
1421 uGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly-AspIleG 1441
6894 -GGACCGTGACAGAAAATTCGCGAGGAGGAACAGCTGCGCCAGGAGGAGGAGGACAGC 6952
1441 lnValGlyMetMet-----AspLysLysGlyGlnLeuGluValGluIleAlaAla 1459
6953 AGCTGAGCGCGCAAGAGCGTGACAGAAAATTCGTTTGAAGGAACAGAAAAGTGGCGCGC 7012
1459 rGlyLeuValLys-----ProGlySerLys 1469
7013 A-GGAACAAGAGAGAAAATTCATGGAGGAGCAACAGCAGCTGCGCGCGCAGGAGGCCAA 7071
1469 hrLeuProAlaProTyrValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLysL 1489
7072 CAACAGCTCGGCCA-----GGAGGACAGAAAATTCGCGGAA 7107

QY 1489 yslYsThrLys-ValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSer--- 1507
DB 7108 GACGAACAGCTGCTCCAGGAAGGAGGAACAGCAGCTGCACCGCAAGAGCGTGACAGA 7167
QY 1508 -----PheGluGluSerProGlnGlyArg 1515
DB 7168 AAATTCCTCGAGGAGGAACCGCAGCTGCGC 7197

RESULT 2

US-08-800-644-93
; Sequence 93, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; NUMBER OF SEQUENCES: 117
; METHOD OF INVENTION: Methods of Using Same
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/056,200
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/POCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1507..1644
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1645..2511
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2512..8070
US-08-800-644-93

Alignment Scores:
Pred. No.: 2,27e-14 Length: 9551
Score: 360.50 Matches: 324
Percent Similarity: 33.02% Conservative: 242
Best Local Similarity: 18.90% Mismatches: 568


```

QY 1309 rGlnSerAspThrAlaValGlyAlaLeuGlyThr-----SerGlyLysLysArgAr 1326
Db 6442 CGCTCGAGGAGCGCGCAACTCGCTCAGGACCGGCACAGAAATTCGCCGAGGAGAA 6501
QY 1326 gSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerAl 1346
Db 6502 CAGCAGCTAGCGCGCAAGAGCGTGACAGAAATTCGTGAAGAGGACACGAGGTGCGC 6561
QY 1346 aSerGln-----LeuSerGlnTh 1352
Db 6562 CGCCAGACAGAGAGAGAAATTCCTGGAGGAGGAACAGCAGCTGCGCCAGGAGCGTAC 6621
QY 1352 rGluGlyGlyLys----- 1358
Db 6622 AGAAATTCGCGAAGAGGACAGCTGCTCCAGAAAGGAGAAACAGCAGCTGCACCGC 6681
QY 1359 -----LeuArgSerThrVa 1363
Db 6682 CAAGAGCGTGACAGAAATTCCTGGAGGAGGAACACAGCTGCGCGCCAGAGCGTGAC 6741
QY 1363 lGlnArgSerThrGluThrGlyLeuAlaVal-----GluMetArgAsnTyrMetThrAr 1381
Db 6742 AGAAATTCGCGAAGAGGACAGCTGCTCCAGAAAGGAGGAACAGCAGCTGCAGGAG 6801
QY 1381 gGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnLe 1401
Db 6802 GAACAGCAGCTGCGCGCCAGCAGCAACCGCAGAGAAATTCCTC-CAGGAGGAACAGCAGCT 6860
QY 1401 uIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLe 1421
Db 6861 GCGCGC-CCAGGAGCGCGGCAACAGCGCGTCA----- 6893
QY 1421 uGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly-AspIleG 1441
Db 6894 -GGACGTCGACAGAAATTCGCGAGGAGGAACAGCTGCGCCAGGAGGAGGAGGACAGC 6952
QY 1441 lNValGlyMetMet-----AspLysLysGlyGlnLeuGluValGluIleIleArgAla 1459
Db 6953 AGCTGAGCGCGCAGAGCGTGACAGAAATTCGTTTAGAGGACAGAAAGTGGCGCGCC 7012
QY 1459 rGlyLeuValValLys-----ProGlySerLysT 1469
Db 7013 A-GGAACAGAGAGAAATTCATGGAGGAGGACAGCAGCTGCGCGCCAGGAGGCGCAA 7071
QY 1469 hLeuProAlaProTyrValLysValTyrLeuLeuAspAenglyValCysIleAlaLysL 1489
Db 7072 CAACAGCTCGGCA-----GGAGGACAGAGAAATTCGCGGAA 7107
QY 1489 yLysThrLys-ValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSer--- 1507
Db 7108 GACGACAGCTGCTCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7167
QY 1508 -----PheGluLysSerProGlnGlyArg 1515
Db 7168 AAATTCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7197

```

RESULT 3

```

US-08-931-999-4
; Sequence 4, Application US/08931999
; Patent No. 6043219
; GENERAL INFORMATION:
; APPLICANT: Iandolo, John J.
; APPLICANT: Crupper, Scott S.
; TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hovey, Williams, Timmons & Collins
; STREET: 2405 Grand Boulevard, Suite 400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: U.S.A.
; ZIP: 64108

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,999
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/710,561
; FILING DATE: 19-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26,262
; REFERENCE/DOCKET NUMBER: 25043-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 816/474-9050
; TELEFAX: 816/474-9057
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6755 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Staphylococcus aureus
; ORIGINAL SOURCE:
; STRAIN: UT0007
; US-08-931-999-4

```

```

Alignment Scores:
Pred. No.: 3,13e-11 Length: 6755
Score: 311.00 Matches: 288
Percent Similarity: 34.27% Conservative: 226
Best Local Similarity: 19.20% Mismatches: 593
Query Match: 3.75% Indels: 399
DB: 3 Gaps: 53

```

US-09-617-099B-1 (1-1590) x US-08-931-999-4 (1-6755)

```

QY 8 ArgGlyArgProAlaProThrProAlaLalaSerGlnProProGln-----ProGlu 25
Db 61 AGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 26 MetProAspLeuSerHisLeuThrGluGluArgLysLysLysLysLysLysLysLys 45
Db 121 CAGCCACACACCGGACACGC-GCGCAGGAGGAGAG-----AGC 158
QY 46 ArgGlnLysLysGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 61
Db 159 AAAGGACACAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 218
QY 62 -----GluGluHisLysLysLysLysLysLysLysLysLysLysLysLysLys 68
Db 219 GAAACGAAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 278
QY 69 ThrGlnTyrPheProPheSerGlyThrGluLeuValAsnValLeuGlnProGln 88
Db 279 GCGAAGGCG-----GGAGACACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 308
QY 89 GlnLysGlnProAsnGluLysGluProGlnThrLysLeuHisGlnGlnPheGluMetTyr 108
Db 309 AAAAACAACACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 368
QY 109 LysGluGlnValLysLysMetGlyGluSerGlnGlnGlnGlnGlnGlnGlnGlnGln 128
Db 369 AGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
QY 129 AlaProThrCysGlyLysCysHisLysThrLysPheAlaAspGlyCysGlyHisAsnCys 148

```

419 ----- 419

L49 SerTyrCysGlnThrLysPheCysAlaArgCysGlyValSerLeuArgSerAsn 168
::: ||||| |
420 AACAGGACACAACGGAA-----GGGGGGGACGAGAAAAAAAAGGACCAA 467

L69 LysValMetTrpValCysAsnLeuCysArgLysGlnGlnGluLeuThrLysSerGly 188
||||| :|||
468 -----ACGAAGAAAAACCAAGAAAAAACAACCGAAGAAGAAAA----- 506

L89 AlaTrpPheTyrAsnSerGlySerAsnThrLeuGlnProAspGlnLysValProArg 208
| | | | |
507 -----GAAAGAACCAAGCCAGAGACCACAAAGG-GNACGC 544

L209 GlyLeuArgAsnGluGluAlaProGlnGluLysAlaLysLeuHisGluGlnProGln 228
::: ||||| :|||
545 -----AAAACGGGCCAGAAAAAGAAAAAGCAA-----AGCCAACACACAG 586

L229 PheGlnGlyAlaProGlyAspLeuSer---ValProAlaValGluLysGlyArgAla--- 246
::: ||||| |
587 GA-GAAGGAACAAAAAAGAAAGCAGACGACAAACACAAAGAGAGAGGGGAGGACCA 645

L247 -----HisGlyLeuThrArgGlnAspThrIleLys 256
| | | | |
646 ACGAAAAAAAAAGAACAAAAAGAAAGCCGGAAGCGAAGAAACAGACGACCAAC 705

L257 AsnGlySerGlyValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSer 276
| | | | |
706 AAAACGACAAAACGGCCACACAGAAAAACACACCCAAAAAAGCGCACAGAAAAAAA 765

L277 ProSerValSerArgAspGlnAsnArg-----ArgTyrGluGlnSerGluGluArg 293
::: ||||| |
766 AAGCAAAAAAGAAAAACGAGGAACAGGAACCCAGCAACAAAAAACAAGCGGAAAAAGAC 825

L294 GluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAspTyr 313
::::: |||||
826 GGAIAAAAAACAAACACACACCCCCAAACAAAGAGGCACAAIAAAAAAGGCAACACCC 885

L314 AlaAspArgArgSerGln--ArgGluProGlnPheTyrGluGluProGlyHisLeuAsn 332
| | | | |
886 AAAAAAAGGCACGCAAGAGAGAAAAAAGGGGAAGGAAAA----- 930

L333 TyrArgAspSerAsnArgArgGlyHisArgHisSerLysGluTyrIleValAspAspGlu 352
::: ||||| :|||
931 -----AACCCAAGGAAGGACCAAAAAAGAGGAAAA----- 963

L353 AspValGluSerArgAspGluTyrGluArgGlnArgGluGluGluTyrGlnAlaArg 372
::: ||||| :|||
964 ---AAAAAIAAACCGGAAAAACAGAAAAAAGGAAAAAACAACAGACAGACCAACAA 1020

L373 TyrArgSerAspProAsnLeuAlaArgTyr-ProValLysProGlnProTyrGluGlu--- 391
||||| :|||
1021 CGAAACACGGACCCGAACCAAAAA-----AAGAAGCCAGGCCGAGGAAAAAGAGA 1074

L392 -----GlnMetArgIleHisAlaGluValSerArgAlaArg 403
||| :||| :|||
1075 CAAGAAAAAAGCACAGAGACACAGGAAGGAGAACCCACAAGAAACAAAAAGCAGA 1134

L404 HisGluArgArgHisSerAspValSerLeuAlaAsnAlaGlu-----LeuGluAsp 420
::: ||||| :|||
1135 GAACAAAAAAGCAGCAGCAACCCCAAGAAAGAAAGAAAAAGAACGACAGACAGAC 1194

L421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
||| :||| :|||
1195 AAAGACAAAGAGAAAAAAGCGGCAAAAGCAGAACGACCAAGAGACGNAACAAAGAAAGG 1254

L441 ArgAlaAlaMetGluAsnGlnArgSerTyr-SerMetGluArgThrArgGluAlaGlnGly 460
||| :||| :|||
1255 AAAAAAGGAC-GAAGAAAAAACAAGACACACCAACAAAGACAGACGAAAGAAA--- 1310

L461 GlnSerSerTyrProGlnArgThr-SerAsnHisSerProProThrProArg-ArgSerPro 480
::: ||||| :|||
1311 GAAGACCGGAAACACGACCAAAACG-----GAAGGACGGCCCAAGAAAGAGGAGGCCCA 1364

Qy	481	IleProLeuAspArgProAspMetArgAlaAspSerLeuArgIysGlnHisLeu	500
Db	1365	GGAGAGCGCAGAAAAGGCGC-----AGAACCGCCCGCCAGCAACCGGAAAAA	1418
Qy	501	AspProSerSerAlaValArgLysThrLysArg-GluLysMetGluThrMetLeuArgAs	520
Db	1419	AAAAGGAAAAGAGCGCAGGAAAGACAAAAGAGGAGGAAACACGCAAGAAACACGCG	1478
Qy	520	nAspSerLeuSerSerAspGlnSerGlnSerValArgProProProProProHisLy	540
Db	1479	GGAAGGAGCGCGAGCAGGCAA---AGGAGCCACCGCGGAAGACACAGAAAACAAGAA	1535
Qy	540	sSerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAl	560
Db	1536	AAAACAGCAAAAGCGAAGAGAGAAAAGAAAAGAAAGAGAGAAAAGCGCAGCAGCAGGA	1595
Qy	560	aSer-----ThrProGluTyrThrSerCysAspValGluLeuGluSe	575
Db	1596	GAGCGGAGCAAAAGAGAAAAAACCAAGAGAACCC-----GA	1631
Qy	575	rGluSerValSerGlnLysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGl	595
Db	1632	CGAAACAAACAAAGAAAGGAAGAGCAGCAAAAAGAAAAGAAAACACCGCGAGCCAGG	1691
Qy	595	yValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGl	615
Db	1692	AAAAGCCAGAGAGCAAAAAAACAACAAAGACAGAAAAAGAAAAAGAAAGAACAAAAA	1751
Qy	615	yHisSerLeuGluLysLeuGluTyrSerGluProGlnLysAspSerGlyValas	635
Db	1752	ACAAACAAACACAGCGCAAAAAGAAAAGCGGAGAGCAAAAAACACCGCGGCAAGGA	1811
Qy	635	pThrCysSerSerThrThrLeuAsnGluGluHisSerHisSerAspLys-----HisPr	653
Db	1812	CCGACACAGCAACGAGGAAAGAAACGGAACAAACAGAGAGAGAAAAAGCAGAAAGAA	1871
Qy	653	oValThrTrpGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLy	673
Db	1872	AACCACGAGCGCGAACCAGGAGGAGGAGAGCAAGAAAGCAAGACGAAAGAGAAACAC	1931
Qy	673	sArgLeuLysAspGlySerValProArgAspSerSerGlyAlaMetLeuGlyLeuLysValva	693
Db	1932	GAAGAAGAAAAGAAAAAACA-----	1950
Qy	693	lGlyGlyLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLysValLysLysGl	713
Db	1950	-----	1950
Qy	713	ySerLeuAlaAspThrValGlyHisLeuArgProGlyAspGluValLeuGluTrpAsnGl	733
Db	1951	-----CGAGCCGCGAAGAAAGAAACAAAGAGGAGGAGAGAG	1982
Qy	733	yArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGluSerLysPr	753
Db	1983	AAACACAGCCCGAGGAAAGACA-----GGAAAAACC	2012
Qy	753	oGluProGlnValGluLeuValValSerArgProIleGlyAspIleProArgIleProAs	773
Db	2013	GGAGAAACAA-----GAAAGCAGA-----GGGAAAGAGGCC-----CCCAA	2048
Qy	773	pSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspAr	793
Db	2049	CGAGGGAAGGCAAGAAAAAAGAGAGGCGCCAC-----	2088
Qy	793	gProSerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGlnPh	813
Db	2089	-----ACAAAGCCAAACAGAAAGCGGAGGCGGAGAAAGAA-----	2121
Qy	813	eLeuSerGlyGlnLeuSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleVa	833
Db	2122	-----CGAAAGAGAGCGCAGCAACA-----	2142

Qy	833	lThrIleLeuGlyAlaIysAspLeuProSerArgGluAspGlyArgProArgAsnProTy	853
Db	2143	-----GAAIAAAAGAAACACGAGGAAGCGAAACAGAAAGCGGAA	2189
Qy	853	rValIysIleTy rPheLeuProAspArgSerAspIysAsnIysArgArgThrIysThrVa	873
Db	2190	GAACAA-----AGGAAGAAAGAAAGAAAGAAAGCGGAAGAACCGAAGCCGGA	2240
Qy	873	lLysLysThrLeuGluProIlystTtPAsnGlnThrPheIleTy rSerProValHisArg	893
Db	2241	AAGCGCAACGAGAGCAAGAAACGAG-----AA	2273
Qy	893	gLuphArgGluArgMetLeuGluIleThrLeuTtPAspGln-AlaArgValargLug	913
Db	2274	GGACACAGAGGAGAAAGAAAGAAAGCAAGAAAGCAAGAAAGCAAGCAAGCAAGAA	2333
Qy	913	luGluSerGluPheLeuGlyGluIleLeuIleGluLeuGluThrAlaLeuLeuAspArg	933
Db	2334	CAAAACGAA-----CAACGCAAAAAAACAACAA	2366
Qy	933	luProHisTtTy rLysLeuGlnThrHisAspValSerSerLeuProLeuProArgPro	953
Db	2367	GGCCAAACAAAGAAAGAAAGCAAGGAGCAACAAACACACACACCAAGCCAC	2426
Qy	953	erProTy rLeuProArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArg	973
Db	2427	AC-----AAAAAAAGACAGCACAAAC-CAAAAC	2455
Qy	973	erLysArgIleSerAspSerGluValSerAspTy rAspCysGluAspGlyValGlyVal	993
Db	2456	AAAGAAACACCAAGAAAGCGCGGAGAGAACGACACAA-----	2499
Qy	993	alSerAspTy rArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSerValProGlu	1013
Db	2500	-----GAAIAAAAGGAGAGGCAACAGCGGCAAAACCGCGCGCCGAGGAGA	2554
Qy	1013	lnValMetSerSerAsnHisCysSerProSerGlySerProHisArgValAspValIleG	1033
Db	2555	AAACGAGAAAGGAGC-----AAACGCAAGAAAGGCCACAG	2593
Qy	1033	lyArgThr-ArgSerTtPProSerProSerAlaProProGlnArgAsnValGluGlnGly	1052
Db	2594	GAGCACAGAGAGGCAAGAGGAAACCAAGCAAAAGAAAGCAACAGCAAGCAAGAA	2653
Qy	1053	HisArgGlyThrArgAlaThrGlyHisTy rAsnThrIleSerArgMetAspArgHisArg	1072
Db	2654	CACAA-----	2659
Qy	1073	ValMetAspAspHisTy rSerSerAspArgAspArgAspCysGluAlaAlaAspArgGln	1092
Db	2660	-----CCAAAGGACAGAGAGACAGAAAGGAGCCAAAG	2695
Qy	1093	Pro-----Ty rHisArgSerArgSerThrGluGln	1102
Db	2696	CCAAAGCAGCAAAAGGAAGAGGGAACAGAAACGACAGAGAGGGCAACAGGACAA	2755
Qy	1103	ArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAspThrAsn	1122
Db	2756	AAACGAAGCAGGAAAGAAAGCAAGAGGGGGGCAAGAACCGCCAAAGCCGAC	2815
Qy	1123	LeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAlaLeu	1142
Db	2816	GCAAACAAAA-----CCAAAGGGAAGACAGGACAGCGCGAGCAGACAGCGGAGGCA	2872
Qy	1143	SerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThr	1159
Db	2873	ACAAACAAACACGAAGCGGCGCAACGAAAGAAAGAAAGCAAGAAAGAAACCGAGGAGA	2932
Qy	1160	-----ProGlyThrGlyArgGlyArgGlnLeuProGlnLeuPro	1173
Db	2933	CGAAAGCAGCAAAAGCAGGAGAGAGAAAGAAAGACAGAAAGAAACCGAGAGCA	2992
Qy	1174	ProLysGlyThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLys	1193

Db	2993	GC	AAAAA	AAAAA	AAAAA	AGAGGGG	GAC	CAAG	CAGGCGG	CAGCGGG	ACACAA	3052																	
Qy	1194	Leu	en	ly	s	Ty	Leu	Ala	Ala	Gly	Ser	Asp	Pro	Arg	-----	Leu	1207												
Db	3053	GG	AAAC	CAAG	ACCC	GAAC	CCCG	GCG	CAAA	AGGGG	GAG	AAAA	CA	AAAAA	CA	AAAAA	3112												
Qy	1208	Glu	-Gln	Asp	Ty	-His	Ser	Ly	s	Ty	r	Arg	Ser	Gly	T	ip	Asp	Pro	Hi	e	Arg	Gly	Ala	Asp	Th	1227			
Db	3113	GAC	ACA	GAC	AC	CA	CA	AG	AAG	CAC	CCCA	AA	AG	CA	AG	AA	AG	CA	AG	AA	AG	AG	GGG	GA	CA	CA	AG	3172	
Qy	1227	r	Val	Ser	Th	ly	s	Ser	Ser	Asp	Ser	Asp	Val	Ser	Ala	Val	Ser	Arg	Thr	Ser	-----	Leu	1247						
Db	3173	AA	ACG	AAAA	AA	CA	AC	CG	CC	CA	AC	AG	-----	CA	AA	CA	AA	AG	CG	CC	CA	AG	CA	AG	CA	AG	CA	AG	3226
Qy	1247	r	Ser	Ala	Ser	Arg	Phe	Ser	Ser	Thr	Ser	Ty	Met	Ser	Val	Gln	Ser	Tlu	Arg	Pro	Arg	G	1267						
Db	3227	CG	AA	AGG	GG	CA	AG	CAG	CA	GC	CA	AA	GA	AA	CA	AA	AG	CG	CA	AC	CA	AA	-----	3284					
Qy	1267	Y	Asn	Arg	Ly	s	I	Le	Ser	Val	Phe	Thr	Ser	Ly	s	Met	Gln	Asn	Arg	Gln	Met	Gly	Val	Ser	G	1287			
Db	3285	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----		
Qy	1287	Y	Ly	Asn	Leu	Th	ly	s	Ser	Thr	Ser	Ile	Ser	Gly	Asp	Met	Cys	Ser	Ileu	Glu	Ly	s	Asn	S	1307				
Db	3290	AG	AG	CA	CC	CA	CG	AA	GG	CG	CG	CA	CA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----		
Qy	1307	p	Gly	Ser	Gln	Ser	Asp	Thr	Ala	Val	Gly	Ala	Leu	Gly	Thr	Ser	Gly	Ly	s	Arg	Ser	-----	1327						
Db	3323	GA	AC	CA	G	CA	GA	AA	AA	AG	CA	CC	AA	AA	CG	AA	-GC	CA	GC	CA	CG	CA	CA	CA	CA	CA	CA	CA	3381
Qy	1327	r	Ser	-----	Ile	Gly	Ala	Ly	s	Met	Val	Ala	Ile	Val	Gly	Leu	Ser	Arg	Ly	s	Ser	Arg	Ser	-----	1345				
Db	3382	AG	GA	AA	AG	GA	GG	CG	CG	CA	CA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----		
Qy	1345	r	Ala	Ser	Gln	Leu	Ser	Gln	Thr	Glu	Gly	Gly	Gly	Ly	Ly	Ly	Ser	Leu	Arg	-----	Ser	Thr	Val	1363					
Db	3421	AG	AA	CA	AA	CG	GA	AAAA	AA	AG	GG	G	AG	AAAA	AG	AAAA	AG	AAAA	AG	AAAA	CA	CA							

RESULT 4
US-08-680-897-1/c
; Sequence 1, Application US/08680897
; Patent No. 6008051
; GENERAL INFORMATION:
; APPLICANT: DasSarma, Shiladitya
; APPLICANT: Halladay, John
; APPLICANT: Ng, Wai-lap
; TITLE OF INVENTION: RECOMBINANT VECTOR AND PROCESS FOR CELL
; TITLE OF INVENTION: FLOTATION
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/680,897
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/271,270
 ; FILING DATE:
 ; APPLICATION NUMBER: 944,581
 ; FILING DATE: 14-SEP-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Digiglio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 8680
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 742-4343
 ; TELEFAX: (516) 742-4366
 ; TELEX: 230 901 SANS UR
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9880 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-680-897-1

Alignment Scores:
 Pred. No.: 9,6e-10 Length: 9880
 Score: 292.50 Matches: 356
 Percent Similarity: 33.48% Conservative: 178
 Best Local Similarity: 22.32% Mismatches: 582
 Query Match: 3.53% Indels: 480
 DB: 77 Gaps: 77

US-09-617-099b-1 (1-1590) x US-08-680-897-1 (1-9880)
 QY 10 ArgProAlaPro---ThrProAlaAla-----SerGlnPro-ProProGlnProGluMe 26
 DB 4797 AGAAGTCTCCAGGATCTCGCGGCTGTTGGGGCAGACCTGAGTCCGGGTACAGTATA 4738
 QY 26 tProAspLeuSerHisLeuThrGluGluGluGluGluGluGluGluGluGluGluGluGlu 46
 DB 4737 CCGCATTTAAATGACCTTGACGCGAA-----GGGTACTTGGAGT 4696
 QY 46 gGlnLysLysGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 66
 DB 4695 ACAGAAATTGAGCAAGCGCAA-----GTGTATCGCTCTCCGATCTCGAAGGAGC 4645
 QY 66 a-----GlnProThrGlnTrpPheProPheSerGlyIleThrGlu 80
 DB 4644 GTTCACCCGAATAGACCATGTTGATCATGTTGCTGCTGTTTTCGCTGCTACTCAAGGC 4585
 QY 80 uValAsnAsnValLeuGlnProGlnGlnLysGln-ProAsnGluLysGluProGlnThrL 100
 DB 4584 GGTATGACCGATTGCAAGCGCTCAATCACAATCCCAAGGAAGAGACCAATGACTG 4525
 QY 100 yLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerG 120
 DB 4524 AGAACCTATACATACGATGATCATCGAACAGGAAGATCTCGAATAGATGTCGAAGGCG 4465
 QY 120 lnglnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysP 140
 DB 4464 TTGCGGAGCGGAACAG-----GTCTATACGCTCGATTACAGACACTCT 4420
 QY 140 heAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysG 160
 DB 4419 CCGCT-----GTGCTCTCTGACATTTGATACGACCGACCCCGACGCGACCG 4375
 QY 160 lyGlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysG 180
 DB 4374 ATGAGGACGTGGAGAGCTCATACAAACGTCTCCAGGAGGTATTGAAGCACGAGAGAAC 4315

QY 180 lnglnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuG 200
 DB 4314 GCACAGTCTCCCATGAGCTTCGGGATGGGTTTC-----AAAAGTCCGCGCAGCTAA 4261
 QY 200 lnglnProAspGlnLysValProArgGlyLeuArgAsnGlu-----GluAlaP 216
 DB 4260 AGGGTGATTTCGGCGGGCGGCGAGTGCATTTCGCGAGTACGCTGAATGACATCGAGGAA 4201
 QY 216 roGlnGluLysLysAlaLysLeuHisGluGlnProGlnPheGlnGlnGlnGlnGlnGln 236
 DB 4200 CCGTCGAACCTTGGCGTGAAGATACTC-----GGTCTCGCGACG 4162
 QY 236 euSerValProAlaValGlu---LysGlyArgAlaHisGlyLeuThrArgGlnAspThrI 255
 DB 4161 ATACAGTCCCTCGAGAAAGAAATCCAGGAAACGTCACGATCACTAGCAGA----- 4110
 QY 255 eLysAsnGlySerGlyValLysHisGlnIleAlaSerAspMetProSerAspArgLysAr 275
 DB 4109 -----TCTGATATCAACGAGACCGAGAACG 4084
 QY 275 gser-----ProSerValSerArgAsp----- 282
 DB 4083 ATCTCTTCACAGACCGCTGATCATATAAGTGTACTCTGCTGCTGCTGCTGCTGCTGCTG 4024
 QY 283 -----GlnAsnArgArgTyrGluGlnSerGluGluArgGluAsp----- 295
 DB 4023 ATGCTTTCGACTCCGCCCATCGACGATGTCGAAGCGGAATACGACGAACTGACGATTCAGT 3964
 QY 296 -----TyrSerGlnTyrValProSerAspGlyThrMetPr 307
 DB 3963 ACACGGGGCGCTGGCGCGGTACAAATTCGTGGACATTCACATCGCGCGGCAACAGCAAC 3904
 QY 307 oArgSerProSerAspTyrAla---AspArgArgSerGlnArgGluProGlnPhe----- 324
 DB 3903 AAGGAGCGCGATATGATCATGATGTCATCATAGCATCTCTCTGCTGAGCCCGTTTTCTCT 3844
 QY 325 -----TyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgGlyHisAr 342
 DB 3843 TGCTGGATATCTCCAGAC-----GATGGCTTCACGAGATGTAGC 3802
 QY 342 gHisSerLysGluTyrIleValAspAsp-GluAspVal---GluSerArgAspGluTyrG 361
 DB 3801 ACACCAA-----ATCGATTGAGAGCATCATCAGGAGAACCACTCTCTGTAGC 3754
 QY 361 lu---ArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuA 380
 DB 3753 AAGTTGGTGAACGATCCGACGAGGAGTACCAAGACAGATGCGTACCGCATGGAGGTCA 3703
 QY 380 laArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValS 400
 DB 3702 AATCGCAACTGGAGATGGCGGAGCGGTCCAAAGAACAGATGCGTACCGCATGGAGGTCA 3643
 QY 400 exArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluA 420
 DB 3642 AGAATAATGGTATCCCGCAGAAA-----ACGACGACGCGTCCGACGACCATCTTCCCA 3589
 QY 420 spSerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluA 440
 DB 3588 GCTCAGCGGCTGCTCGACCAATTGCACACGCTCGTTCGAGGTGCTCGCGCATCGAAGA 3529
 QY 440 rgArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnG 460
 DB 3528 GGAGGCGGTGCTACCGACAGCATGAGCGGTATCGACCGGGGAAACCGCCGAATGTATTA 3469
 QY 460 lyGlnSerSerTyrProGln-----ArgThrSerAsnH 471
 DB 3468 CGACTACGACGTTTCCATTCGGGCTGGACGACGCGGCTGCTCTACGACGAGGAACC 3409
 QY 471 iSerPro-----ProThrProArgSerProIleProLeuAspArgProAspMetA 489
 DB 3408 ATCGTCCCAATCGTCCCGCTCCGACAGCGATCCCGTTCAGCAAGAGCA-----C 3358
 QY 489 rgArgAlaAspSerLeuArgLysGlnHisLeuAspProSerSerAlaValArgLysTr 509

Db	3357	 GGAGG- :::	hrlYsArgGluLysMetGluThrMetLeuAAsnAspSerLeuSer- :::	ATTGATTCATCA 3343
Qy	509	 hrlYsArgGluLysMetGluThrMetLeuAAsnAspSerLeuSer- :::	hrlYsArgGluLysMetGluThrMetLeuAAsnAspSerLeuSer- :::	524
Db	3342	CGTCGAGACTCGTGGAGGAAGACGGCGGCGAGCACTCGTCGTCTCGCAGATTGGCCGG 3283	CGTCGAGACTCGTGGAGGAAGACGGCGGCGAGCACTCGTCGTCTCGCAGATTGGCCGG 3283	
Qy	525	-----SerAspGlnSerGluSerValArgProProProArgProHisL 540	-----SerAspGlnSerGluSerValArgProProProArgProHisL 540	
Db	3282	CGTGACGGACGACGACGTCGACGTCGCACTCGATCGATACCG- :::	CGTGACGGACGACGACGTCGACGTCGCACTCGATCGATACCG- :::	3245
Qy	540	ysSerLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuA 560	ysSerLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuA 560	
Db	3244	-----ACAGAGCGGCACTCACGCTACG 3223	-----ACAGAGCGGCACTCACGCTACG 3223	
Qy	560	laSerThrProGluTyrThrSerCysAspAspValGluLeuLysSerGluSerValSerG 580	laSerThrProGluTyrThrSerCysAspAspValGluLeuLysSerGluSerValSerG 580	
Db	3222	CGTCGACGACGACGTCGTGAGC- :::	CGTCGACGACGACGTCGTGAGC- :::	3181
Qy	580	lu-lysGly-----AspSerGlnLys--GlyLysArgLysThrSerGluGln 594	lu-lysGly-----AspSerGlnLys--GlyLysArgLysThrSerGluGln 594	
Db	3180	GGAGCTGGCGATCACCGATATGACTCTCAGAAATCAGGTTCTCGAAATTCGGCTAGCAGC 3121	GGAGCTGGCGATCACCGATATGACTCTCAGAAATCAGGTTCTCGAAATTCGGCTAGCAGC 3121	
Qy	595	GlyVal-----LeuSerAspSerAenThrArgSerGluArg 606	GlyVal-----LeuSerAspSerAenThrArgSerGluArg 606	
Db	3120	AACATCGGACACCGATGAGGTGAATCCATGATGCGACAAACACACGCAAAACACACAG 3061	AACATCGGACACCGATGAGGTGAATCCATGATGCGACAAACACACGCAAAACACACAG 3061	
Qy	607	GlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGluAsp- :::	GlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGluAsp- :::	621
Db	3060	CAGAAAGCACGACGACGACGAGTAAGGCGCAGATCAACCGGATAAAGCCCGGCGCAAT 3001	CAGAAAGCACGACGACGACGAGTAAGGCGCAGATCAACCGGATAAAGCCCGGCGCAAT 3001	
Qy	622	-----LeuGluTrpSerGlu 626	-----LeuGluTrpSerGlu 626	
Db	3000	CTGCTTCGGCAGCGCAAAACTCGCCGCGTCGACCCGCAACCGCAACAGTCGAA 2941	CTGCTTCGGCAGCGCAAAACTCGCCGCGTCGACCCGCAACCGCAACAGTCGAA 2941	
Qy	627	ProGlnLysAspSerGlyValAspThrCysSerSerThrThrLeuAsnGluGluHis 646	ProGlnLysAspSerGlyValAspThrCysSerSerThrThrLeuAsnGluGluHis 646	
Db	2940	CCCCGGGTGACGATTCGACGGCGAGCATACTGACGGGGAGACGGTCACAGAAATCCACG 2881	CCCCGGGTGACGATTCGACGGCGAGCATACTGACGGGGAGACGGTCACAGAAATCCACG 2881	
Qy	647	SerHisSerAspLysHisProValThrTriPdnProSerLysAspGlyAspArgLeuIle 666	SerHisSerAspLysHisProValThrTriPdnProSerLysAspGlyAspArgLeuIle 666	
Db	2880	GCTCACTCG-----ACATGCCGCTCAGAAAGTCGACCGGAG----- 2842	GCTCACTCG-----ACATGCCGCTCAGAAAGTCGACCGGAG----- 2842	
Qy	667	GlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSer----- 684	GlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSer----- 684	
Db	2841	---AACGGGTACGCAACTCCCACTCGACCGTCCGAGAGACTCCGAAAGTACTCGGATGTG 2785	---AACGGGTACGCAACTCCCACTCGACCGTCCGAGAGACTCCGAAAGTACTCGGATGTG 2785	
Qy	685	-----GlyAlaMetLeuLysValValGlyGlyLysMet 697	-----GlyAlaMetLeuLysValValGlyGlyLysMet 697	
Db	2784	ACGGCTGAGACGATGTCGCGCAACGCGCTCCACCAGAAAAACAGCCGCTGGAAATCTG 2725	ACGGCTGAGACGATGTCGCGCAACGCGCTCCACCAGAAAAACAGCCGCTGGAAATCTG 2725	
Qy	698	ThrGluSerGlyArgLeuLysAlaPheIleThrLysValLysLysGlySerLeuAlaAsp 717	ThrGluSerGlyArgLeuLysAlaPheIleThrLysValLysLysGlySerLeuAlaAsp 717	
Db	2724	GAGGCATCCGCC----- 2713	GAGGCATCCGCC----- 2713	
Qy	718	ThrValGlyHisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGln 737	ThrValGlyHisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGln 737	
Db	2712	-----CGTCGGATGAG-----ACGCGCTCC 2692	-----CGTCGGATGAG-----ACGCGCTCC 2692	
Qy	738	GlyAlaThrPheGluGluValTyrAsnIleIleLeuLysSerLysProGluProGlnVal 757	GlyAlaThrPheGluGluValTyrAsnIleIleLeuLysSerLysProGluProGlnVal 757	
Db	2691	GGAGTCCCACTGAGGACGAGGTGAAC-----GATGAGTACCCCAACC 2647	GGAGTCCCACTGAGGACGAGGTGAAC-----GATGAGTACCCCAACC 2647	
Qy	758	GluLeuValValSerArgPro-----IleGlyAspIleProArgGlyProAspSer 774	GluLeuValValSerArgPro-----IleGlyAspIleProArgGlyProAspSer 774	
Db	2646	GACCGCTCGCAGGCGGACCTCGCCGAGATGCTGGAGATGCTCTAGACA-AAGGATGTCG 2588	GACCGCTCGCAGGCGGACCTCGCCGAGATGCTGGAGATGCTCTAGACA-AAGGATGTCG 2588	
Qy	775	ThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArgPro 794	ThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArgPro 794	
Db	2587	TCGTCAACGGGATATCCGAGTCAGCTCAGCGTCGGGACACGGAAC----- 2546	TCGTCAACGGGATATCCGAGTCAGCTCAGCGTCGGGACACGGAAC----- 2546	
Qy	795	SerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeu 814	SerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeu 814	

2545	-----	TCCTCGGTATCAGTTACGGCCCGCGATGTGCTTCGTTCGACACACAGCGCCGAAT	2492	
815	SerGlyGlnLeuSerIle	-----LysLeuTirPhe	-----AspLysValGly	828
2491	ACGGCTCGAGTTCCCAACGGGACGGATATGGAGCGGTCGAGTCGCGCGGATATCT	2432	ACGGCTCGAGTTCCCAACGGGACGGATATGGAGCGGTCGAGTCGCGCGGATATCT	2432
829	HisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArgGluAspGlyArg	848	HisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArgGluAspGlyArg	848
2431	CACCGGACCACTCGCAC	2411	CACCGGACCACTCGCAC	2411
849	ProArgAsnProTyrValLysIleTyrPheLeuProAspArgSerAspLysAsnLysArg	868	ProArgAsnProTyrValLysIleTyrPheLeuProAspArgSerAspLysAsnLysArg	868
2410	CGTCGAAA	2375	CGTCGAAA	2375
869	ArgThrLysThrValLysLysThrLeuGluProLysTirPAsnGlnThrPheIleTyrSer	888	ArgThrLysThrValLysLysThrLeuGluProLysTirPAsnGlnThrPheIleTyrSer	888
2374	CACGTGCGGACGATTTCGACGCCGACGCATCCA	2342	CACGTGCGGACGATTTCGACGCCGACGCATCCA	2342
889	ProValHisArgGluGluPheArgGluArgMetLeuGluIleThrLeuTirPAspGlnAla	908	ProValHisArgGluGluPheArgGluArgMetLeuGluIleThrLeuTirPAspGlnAla	908
2341	-----CGTCACGCGAGGACGACCAAAATCAGATGCAATAGCCTCGACGACGACGCG	2291	-----CGTCACGCGAGGACGACCAAAATCAGATGCAATAGCCTCGACGACGACGCG	2291
909	ArgValArgGluGluGluSerGluPheLeuGlyCyluIleLeuIleGluLeuGluThrAla	928	ArgValArgGluGluGluSerGluPheLeuGlyCyluIleLeuIleGluLeuGluThrAla	928
2290	GACGATTTCGACGCGCGGCTCACCGCGCTGTGTACTCTGTCTGCAATTCGTCTGTCGTCGAG	2231	GACGATTTCGACGCGCGGCTCACCGCGCTGTGTACTCTGTCTGCAATTCGTCTGTCGTCGAG	2231
929	LeuLeuAspAspGluProHisTirTyrLysLeuGlnThrHisAspValSerSerLeuPro	948	LeuLeuAspAspGluProHisTirTyrLysLeuGlnThrHisAspValSerSerLeuPro	948
2230	GCCCTCGAACAAAGAA	2186	GCCCTCGAACAAAGAA	2186
949	LeuProArgProSerProTyrLeuProArgArgGlnLeuHisGlyGluSerProThrArg	968	LeuProArgProSerProTyrLeuProArgArgGlnLeuHisGlyGluSerProThrArg	968
2185	-----	2168	-----	2168
969	ArgLeuGlnArgSerLysArgIleSerAspSerGluValSerAspTyrAspCysGluAsp	988	ArgLeuGlnArgSerLysArgIleSerAspSerGluValSerAspTyrAspCysGluAsp	988
2167	CGATTGGGGCGACAAATTACAGGCCCTCGAAGACGAGCTCGACGACGTAAACACGACGAG	2108	CGATTGGGGCGACAAATTACAGGCCCTCGAAGACGAGCTCGACGACGTAAACACGACGAG	2108
989	GlyValGlyVal	1007	GlyValGlyVal	1007
2107	GATATCAACGCCGAGGTCTCGAGTTAGG	2057	GATATCAACGCCGAGGTCTCGAGTTAGG	2057
1008	LeuSerValProGluGlnValMetSerSerAsnHisCysSerProSer	1025	LeuSerValProGluGlnValMetSerSerAsnHisCysSerProSer	1025
2056	CGCGACGCCATCGAACAGCTA	2003	CGCGACGCCATCGAACAGCTA	2003
1026	Pro	1040	Pro	1040
2002	CCGGAATCCAGCGTGATGACGTATGACTACACCGCGCCACCCCGGAGAGAGAG	1946	CCGGAATCCAGCGTGATGACGTATGACTACACCGCGCCACCCCGGAGAGAGAG	1946
1041	SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly	1060	SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly	1060
1945	-----AGACCACACGCAACG	1907	-----AGACCACACGCAACG	1907
1061	HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer	1080	HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer	1080
1906	GCTATCTATCTACTCGGTGTCATACCAACGCTCGTCCG	1871	GCTATCTATCTACTCGGTGTCATACCAACGCTCGTCCG	1871
1081	AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr	1100	AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr	1100
1870	-----	1841	-----	1841
1101	GluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAsp	1120	GluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAsp	1120
1840	-----	1811	-----	1811
1121	ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro	1140	ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro	1140
1810	-----	1772	-----	1772
1141	AlaLeuSerArg	1160	AlaLeuSerArg	1160
1771	-----	1724	-----	1724

QY 1160 oGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
Db 1723 AGCAGTGTGTCACCGCGC----- 1705
QY 1180 gSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsn-LysTyrLysGlnV 1200
Db 1704 -----GAGCGAGCGTTCGGTACCGCGTCCGATCGGATTCGACCGG 1661
QY 1200 al---AlaGlySerAspProArgLeuGlu-----GlnAspTyrHisSerLysTyrA 1216
Db 1660 TCCTCGAGCGCGTGTATCGAGTATCGAAGCGTGTGTAGAACACCACTACGAGGCTTCC 1601
QY 1216 rGSerGly-----TTPAsp- 1220
Db 1600 GCGAGCAATTACGTCGTCGCGGAGTGTGGAGTATCGAATCAATCTGTGTGGATT 1541
QY 1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerValSerVal 1240
Db 1540 CCGCAGCGTTCGAGGAGCATTCGAGACCGGAGCGCGCTCCGAGAACCTACGACAGC 1481
QY 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSer-SerThrSerTyrMetSerVa 1260
Db 1480 GCGAGCAACAATCGCGCGCAGGAGAAATCTCTCTCTGAGAAACA-----GT 1433
QY 1260 lGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAs 1280
Db 1432 CCGATCAGCGATTCGAAAGAGTGAACGAGAGCGCGGAGCGGAACTAGCAGATCACTGA 1373
QY 1280 nArg-GlnMetGlyValSerGlyLysAsnLeuThrLys-----SerThrSerLys 1297
Db 1372 AGAGCGCATTACCGCGTCTGTCAGCAGCTGACCGACAGCAGCAGATACCGCGCTAC 1313
QY 1297 exGlyAspMetCysSerLeuGluLysAsnAsp-----Glys 1309
Db 1312 AGGACGACACATCGTCTCATCGAAGAAAGACAGATCGTGGCGTTCGCGGAGC 1253
QY 1309 exGlnSerAspThrAlaValGly---AlaLeuGlyThr-SerGlyLysLysArg----- 1325
Db 1252 AGGACGACAGACCGCTCTCGGTGATCGATTGGATACGATCTCGAACACGAGGAGGTAG 1193
QY 1326 ArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSer 1345
Db 1192 AGATCAGAT-----TCAGCGCGCGTGGCCACCGTACA 1160
QY 1346 AlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArg 1365
Db 1159 CGTTCGCGCCAGATATTGGTAAATATGAGGACCAACAAAGACGACACACGCGATCT 1100
QY 1366 SerThrGluThrGlyLeuAlaValGluMetArgAsnTyrMetThrArgGlnAlaSerArg 1385
Db 1099 -----TGAGTTCGTCGACGCTTACTGCGCGAGCGGCGCT 1064
QY 1386 GluSerThr-----AspGlySerMetAsnSerTyrSerSerGluGlyAsnLeu 1401
Db 1063 GATTACGCGGCGGATGATCGTACGCTGCCGACATTCCTTCATCGGATCAGCCT--- 1007
QY 1402 IlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLeu 1421
Db 1006 -----CCGGC-----AGCGATTG---TGGCATGACCCATCAGCGAGTA 968
QY 1422 GlyProAlaGlnLeuValGly 1428
Db 967 CGGCTGTTCGAGGAGTGGGA 947

RESULT 5

US-08-961-527-363
; Sequence 363, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 363:

SEQUENCE CHARACTERISTICS:

LENGTH: 4483 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-363

Alignment Scores:

Pred. No.: 4,62e-10 Length: 4483
Score: 290.00 Matches: 295
Percent Similarity: 33.68% Conservative: 221
Best Local Similarity: 19.26% Mismatches: 632
Query Match: 3.50% Indels: 388
DB: 4 Gaps: 61

US-09-617-099B-1 (1-1590) x US-08-961-527-363 (1-4483)

QY 7 ProArgGlyArgProAlaPro-----ThrProAlaAlaSerGlnProPro 22
Db 324 COTCAGCATCGACAGCGCTCAGCTTCAGCAGTACCGAGTCTTCAGCCTCAGCGTCCA 383
QY 23 GlnProGluMetProAspLeuSerHisLeuThrGluGluGlu-----Arg 37
Db 384 CAGTGGCTCGCGCTCAACCGATGATCTGATCGGCATCAACCGTGGTGGTGGTGGT 443
QY 38 LysIleIleLeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerVal 57
Db 444 CAAGTACTAGTGCATCAGCTTCAGCATCAACGAGTGCATCGCTTCAGCATCAACGAGTG 503
QY 58 LeuLysIleLysGluGluHisLysAlaGlnProThrGlnThrPheProPheSerGlyIle 77
Db 504 COTCGCTTCAGCGTCAACCGATGCTCAG-----CTTCAGCAAGTACCGATG 554
QY 78 ThrGluLeuValAsnAsnValLeuGlnProGlnLysGln---ProAsnGluLysGlu 96
Db 555 CAGTCTCAGCATCAACAGTCTTCAGCTTCAGCATCAACGAGTGGCTTCGGCTTCAGGCA 614
QY 97 ProGlnThrLysLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGly 116
Db 615 GCACATCAGCATCTGATCAGCGTCAACGCGCTCAGCTTCAGCAAGTACCGATGCTGGT 674
QY 117 GluGluSerGlnGlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHis 136
Db 675 CAGCGCTCAGCGTCAACAGTGGTGGT-----CCTCAGCAAGTACTAGTGGAT 722
QY 137 LysThrLys-PheAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCys 156

```

723 CAGCTTCAGCATCAACGAGTGCATCGCTTCGGCGTCAACGAGTGCATCAGAGTCAGCAA 782
156 sAlaArgCysGlyGlyArgValSerLeuArgSerAsnLysValMet---TrpValCysAs 175
783 GTACAGTGC-----GTCAGCTTCGGCATCAACAGTGCCTCGGCTTCAGCAAGCA 833
175 nLeuCysArgLysGlnGlnGluLeuThrLysSerGlyAlaTrpPheTyrAsnSerG1 195
834 CCAGTGCCT-CGGCTTCAGCAAGTACTAGCGCTT-----CAGCCTCA 874
195 ySerAsnThrLeuGlnGlnProAspGlnLysValProArgGlyLeuArgAsnGluGlu-- 214
875 GCCTCAACCAAGTGCCTCAGCGCTCAGCAAGTATCTCAGCGCTGTAATCGGCATCAACGAGT 934
215 -----AlaProGlnGlnLysLysAlaLysLeuHisGlu----- 225
935 GCGTCCGCTTCAGCAAGTACTAGCGCTCAGCCTCAGCGTCAACAGTGCATCGGCTTCA 994
225 ----- 225
995 GCGTCAACGAGTGCCTCTGAATCGGCATCAACGAGTGCCTTCAGCAAGTACTAGC 1054
226 ---GlnProGlnPheGlnGlyAlaPro-----G1 234
1055 GCCTCAGCCTCAGCGTCAACAGTGCATCGGCTTCAGCATCAACGAGTGCCTCGGCTTCA 1114
234 yAspLeuSerValProAlaValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspTh 254
1115 GCAAGTACTAGCGCTCAGCCTCAGCGTCAACAGTGCATCGGCTTCAGCGTCAACGAGT 1174
254 r---IleLeuAsnGlySerGlyValLys-----HisGlnIleAlaSerAs 268
1175 GCGTCTGAGTCAGCATCAACGAGTGCCTCAGCCTCAGCAAGCATCAGCTTCT-GAATC 1233
268 pMetProSerAspArgLysArgSerProSerValSerArgAspGlnAsnArgTyrG1 288
1234 TGCATCAACAGCAGTGCCTCAGCCTCAGCATCGCAAGCGCCTCAGCTTCAGCAAGTACCAG 1293
288 uGlnSerGluGluArgGluAspTyrSer-----GlnTyrValProSerAspGlyThrMe 306
1294 TGGTTCAGCTCAGCGTGCAGCAAGTGCCTCGCTTCAGCAAGTACCAGTGC-----CGTCA 1347
306 tProArgSer---ProSerAspTyrAlaAspArgArgSerGlnArgGluProGlnPheTy 325
1348 GCCTCAGCAAGTACCAGTGCCTCAGCCTCAGCGTTCAGCAAGTGCCTCGGCTCAA----- 1402
325 rGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGlyHisArgHisSerLy 345
1403 -----CCAGTGCATCTGAATCGGCATCAACAGTGCCTCAG----- 1438
345 sGluTyrIleValAspAspGluAspValGluSerArgAspGluTyrGluArgGlnArgAr 365
1439 -----CCTCAGCAAGTACTAGCGCTCAGCCTCAGCATCAACAGTGCCTCGGCTCC 1485
365 sGluGluGluTyr-----GlnAlaArgTyrAr 374
1486 GCTTCAGCAAGTACTAGTGCATCAGCTTCAGCAAGTACTAGCGCTCAGCTCAGCGTGC 1545
374 gSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMetAr 394
1546 ACAAGCGCTTCAGCTTCAGCAAGTACCAGTGCCT---CAGCCTCAGCGTTCGACAGTGC 1602
394 g-----IleHisAlaGluValSerArgAlaArgHI 404
1603 TCGGCTTCAGCAAGTACTCAGCGTCTGAATCAGCATCAACAGTGCCTCGGCTTCAGCA 1662
404 sGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAspSerArgIleSe 424
1663 TCACAGTGCAT-----CAGCTTCAGCATCAACAGTGCCT 1698
424 rLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArgArgAlaAlaMe 444

```

```

1699 TCAGCTT-----CAGCAAGTACCAGTGCCTCGG----- 1726
444 tGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSerTy 464
1727 -----CTTCAGCATCAACAGTGCCTTCAGTCTCAG----- 1756
464 rProGlnArgThrSerAsnHisSerProProThrProArgArgSerProIleProLeuAs 484
1757 -----CGTCAACAGTGCCTCTGAA----- 1776
484 pArg-----ProAspMetArgArgAlaAspSerLeuArgLysGlnHisLe 500
1777 TCGGCATCAACAAGTGCCTCGGCTTCAGCAAGCAGCAGTGCCTTCGGCTTCAGCGT----- 1831
500 uAspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAs 520
1832 -----CAACGAGTGCCTCTGAGTCAGCATCAACGAGTGCCTCAGCAGCAAGCA 1881
520 n-----AspSerLeuSerSerAspGlnSerG1 529
1882 CATCAGCTTCTGAATCTGCATCAACGAGTGCCTCAGTTCGGCATCAACAGCGCTTCGGC 1941
529 uSerValArgProProProArgProHisLysSerLysLysGlyGlyLysMetArgG1 549
1942 CTCAGCA-----AGTACAAGTGCCTCAGCCTCAGCATCAACAGTGCATC 1986
549 nValSerLeuSerSerSerGluGluLeuAlaSerThrProGluTyrThrSerCysAs 569
1987 AGCTTCAGCCTCAACAGTGCCTCAGCCTCAGCGTCAACAGTGCCTTCGCGCTTCAGCAAG 2046
569 pAspValGluLeuGluSerGluSerValSerGluLysGlyAspSerGlnLysGlyLysAr 589
2047 TACCAGTGCCTCAGCTTCAGCAAGTGCCTCAGCTTCAGCATCAACAGTGCCTTC 2106
589 gLysThrSerGluGlnGlyValLeuSerAspSerAsnThrArgSerGluArgGlnLysLy 609
2107 GCGTTCGGCATCAACAAGTGCCTCAGCATCAGCATCAACAGTGCCTCAGTGCCTCAGCAAGT 2166
609 sArgMetTyrTyrGlyGlyHisSerLeuGluGluAspLeuGluTyrSerGluProGlnI1 629
2167 ACTAGTGCATCAGCATCAGCT-----CAACAGTGC 2199
629 eLysAspSerGlyValAspThrCysSerSerThrThrLeuAsn-----G1 644
2200 TCAGCCTCAGCAAGTA-----TCTCAGCTCTGAATCGGCATCAACAGTGC 2247
644 uGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAspAr 664
2248 TCAGCATCAGCATCAACAGTGCATCGGCTTCAGCGTCAACAGTGCATCAGTGCCTCAGCA 2307
664 gLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSe 684
2308 AGCACCAGTGCCT-----CGGCTTCAGCATCAACAGTGCCTCAGCCT-- 2350
684 rGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSerGlyArgLeuCy 704
2351 -----CAGCAAGTACTCAGCTCTGAATCGGCATCAACAG----- 2386
704 sAlaPheIleThrLysValLysGlySerLeuAlaAspThrValGlyHisLeuArgPr 724
2387 -----GTGCGTCA 2394
724 oGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluVa 744
2395 GCGTTCAGCAAGTA-----CTAGTGCATCAGCATCAGCATCAACAGCT 2436
744 lTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuValValSerArgPr 764
2437 GCATCG-----CTTCAGCAAGTACCAGCGCTCAGCTTCAG-----CAAGCACC 2481
764 oIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGluSerSerSerSe 784
2482 AGTGCCTCAGCTCAGCAAGTACCA----- 2506

```

QY 784 rSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSerProMetSerPr 804
DB : : : : :
DB 2507 -GCGCTCAGCTCAGCAAGCAGCAGTGGCTCAGCTTACGCAAGTACGAGTGGTTCAGCC 2565
QY 804 o-----GlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSe 819
DB : : : : :
DB 2566 TCAGCGTGACAAAGTGGCTGGCTTCAGCAAGTACTCAGCTCTGAATCAGCATCAAG 2625
QY 819 rIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAla-- 838
DB : : : : :
DB 2626 AGTG-----CATCAGCTTCAGCATCAACAAGTGGCTTCAGCTTCAACCAAGTACC 2673
QY 839 -----LysAspLeuProSerArgGluAspGlyArgProArgAs 851
DB : : : : :
DB 2674 AGTGGTGGCTTCAGCATCAACGAGTGGCTTCAGCTTCAGCGTCAACCAAGTGGCTTGA 2733
QY 851 nProTyrValLysIleTyrPheLeuProAspArgSerAspLysAsnLysAlaArgThrIly 871
DB : : : : :
DB 2734 TCAGCATCAACAAAGTG-----CCTCGGCTTCAGCAAGCAGCAGTGGCTTCAGCTCA 2784
QY 871 sThrValLysLysThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHi 891
DB : : : : :
DB 2785 G-----CAAGTACTAGTGCATGGCTTCAGCA 2811
QY 891 sArgArgGluPheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValAr 911
DB : : : : :
DB 2812 TCAGCAAGTGGCTTGAATCGGCATCAACGAGTGGCTTCAGCTTCAGCATCAACGAGTGG 2871
QY 911 gGluGluGluSerGluPheLeuGlyGluIleLeuIleGluLeuGluThrAlaLeuLeuAs 931
DB : : : : :
DB 2872 TCAGCGCTCAG----- 2881
QY 931 pAspGluProHisTrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeuPro-- 950
DB : : : : :
DB 2882 -----CAAGCATCATCAGCTTCTGAATCTGCATCAACCACT 2916
QY 951 -ArgProSerProTyrLeuProArgArgGlnLeuHisGlyGluSerProThrArg---Ar 969
DB : : : : :
DB 2917 GCGTGGCTTCAGCGTCAACCAAGTGGCTGGCTTCAGCTCAGCAAGTGGCTTCAGCTTCA 2976
QY 969 gLeuGlnArgSerLysArgIleSerAspSerGluValSerAspTyrAspCysGluAspGl 989
DB : : : : :
DB 2977 GCATCAACAGTGGCTGGCTTCAGCAAGCAGCAAGTACCTCAGCGTCACTCCGCT-- 3034
QY 989 yValGlyValValSerAspTyrArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSe 1009
DB : : : : :
DB 3035 -----CAACAGTGGCTTCGCTTCAGCAAGCAGCAAGTGG 3069
QY 1009 rValProGluGlnValMetSerSer---AsnHisCysSerProSerGlySerProHisAr 1028
DB : : : : :
DB 3070 TCAGCGCTCAGCAAGTATCTCAGCGTCTGAATCGGCATCAACGAGTGGCTTCAGCA 3129
QY 1028 gValAspValIleGlyArgThrArgSerTrpSerProSerAlaProProGlnArgAs 1048
DB : : : : :
DB 3130 AGCGCAAGTACTCAGCGT-----CAGCTTCGCGCTCAACCAAGTGGCTG 3174
QY 1048 nValGluGlnGly-----HisAr 1054
DB : : : : :
DB 3175 GCTTCAGCAAGCAAGTGGCTCAGCGTTCAGCAAGTATCTCAGCTCTGAATCGGCATCA 3234
QY 1054 gGlyThrArgAlaThrGlyHisTyrAsnThrIleSerArgMetAspArgHisArgValMe 1074
DB : : : : :
DB 3235 ACAGTGGCTCAGTGGCTCAGCATCAACGAGTACGCTCAGCGCTCAGCAAGCAGCATCAGCTTCT 3294
QY 1074 tAspAspHisTyrSerSerAspArgAspCysGluAlaAlaAspArgGlnProTy 1094
DB : : : : :
DB 3295 GAATCGGCATCAACCAAGTG-----CGTCAGCGCTCA 3324
QY 1094 rHisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThrArgSerAr 1114
DB : : : : :
DB 3325 GCATCGA-----CAAGCGCTCAGCTTCAGCA-AGTACCAGTGGCTTCAGC 3368

QY 1114 gSerSerGluArgProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyAr 1134
DB : : : : :
DB 3369 CTCAGCGTGCACAAGTGGCTCAGCTCAACAGTGGCTCTGAATCGGCATCAACAGTGC 3428
QY 1134 gSerAlaProProSerProAlaLeuSerArgSer-HisProArgThrGlySerValGlnT 1154
DB : : : : :
DB 3429 CTCAGCGTTCAGCAAGTACTAGTGCATCAGCTTCAGCATCAACGAGTGGCTTCAGC 3488
QY 1154 hrSerProSerSerThrProGlyThrGlyArgArgGly---ArgGlnLeuProGlnLeuP 1173
DB : : : : :
DB 3489 ATCAACCA-----GTGCTCGGCTTCAGCGTCAACCAAGTGGCTTCAGCTTCAGCAAGTAC 3542
QY 1173 roProLysGlyThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetL 1193
DB : : : : :
DB 3543 CAGTG-----CTTCAGCTTCAGCATCAACAGTGC 3572
QY 1193 ysLeuAsnLysTyrLysGlnValAlaGlySerAspProArgLeuGluGlnAspTyrHis 1213
DB : : : : :
DB 3573 TTCAGCGCTCAGCATCGACAAGTG-----CCTCGGCTTCAGCAAGCAGCATCAGC 3620
QY 1213 erLysTyrArgSerGlyTrpAspProHisArgGlyAlaAspThrValSerThrLysSers 1233
DB : : : : :
DB 3621 ATCTGAATCAGCGTGCACAAGCGCTCAGCT-TCAGCAAGTACCAAGTGGCTTCAGCCTCAG 3679
QY 1233 erAspSerAspValSerAspValSerAlaValSerArgThrSerSerAlaSerArgPheS 1253
DB : : : : :
DB 3680 CGTCAGCAAGTGGCTCAGCTCAGCATCAGTACTAGTGCATCAGCTTCAGCATCAACAGTGC 3739
QY 1253 erSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleSerAr 1273
DB : : : : :
DB 3740 CATCGGCTTCGCGTCAACCAAGTGCATCAGATCAGCAAGTACCAAGTGGCTTCAGCTTCOG 3799
QY 1273 alPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLys- 1292
DB : : : : :
DB 3800 CATCAACAAGTGGCTTCGCGTTCAGCAAGCAGCAGTGGCTTCGCTTCAGCAAGTACTAGCG 3859
QY 1293 -----SerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspG 1308
DB : : : : :
DB 3860 CTCAGCGCTCAGCGCTCAACCAAGTGGCTCA-----GCCTCAGCAAGTATCTCAGCGT 3910
QY 1308 lySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSers 1328
DB : : : : :
DB 3911 CTGAATCGGCATCAACGAGTGGCTTCGCTTCAGCAAGTACTAGCGCTTCAGCGCTCAGCGT 3970
QY 1328 erIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerG 1348
DB : : : : :
DB 3971 CAACAAGTGCATCGGCTTCAGCGTCAACGAGTGGCTTCGAATCGGCATCAACGAGTGGCT 4030
QY 1348 nLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerThrG 1368
DB : : : : :
DB 4031 CGGCTTCAGCAAGTACTAGCGCTCAGCGCTCAGCGTCAACAGTGCATCGGCTTCAGCAT 4090
QY 1368 luThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerT 1388
DB : : : : :
DB 4091 CAACAGTGGCTTCGCGTTCAGCAAGTACTAGCGCTTCAGCGCTCAGCGTCA---ACAAGTG 4147
QY 1388 hrAspGlySerMetAsnSerTyrSerSerGlu 1398
DB : : : : :
DB 4148 CATCGGCTTCAGCGTCAACGAGTGGCTTCGAG 4179

RESULT 6

US-09-144-085-3/c

; Sequence 3, Application US/09144085

; Patent No. 6280999

; GENERAL INFORMATION:

; APPLICANT: Gustafsson, Claes

; APPLICANT: Betlach, Mary C.

; APPLICANT: Ashley, Gary

; APPLICANT: Julien, Bryan

; APPLICANT: Ziermann, Rainer

; TITLE OF INVENTION: SOPANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 30062-20020.20

```

; CURRENT APPLICATION NUMBER: US/09/144,085
; CURRENT FILING DATE: 1998-08-31
; EARLIER APPLICATION NUMBER: 09/010,809
; EARLIER FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 33529
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-144-085-3

```

Alignment Scores:	
Pred. No.:	1.01e-08
Score:	289.50
Percent Similarity:	31.55%
Best Local Similarity:	21.40%
Query Match:	3.48%
DB:	3
Length:	33529
Matches:	348
Conservative:	165
Mismatches:	581
Indels:	534
Gaps:	72

US-09-617-099B-1 (1-1590) x US-09-144-085-3 (1-33529)

QY	2	SerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnProPro	21
Db	15682	TCAGGCCACGGGGCAAGC---CAGCCCGCTCCAGTCTCCACGCCCTCTCCAGCA	15626
QY	22	Pro-----GlnProGluMetProAspleuSerHisLeuThr-GluGluGluAr	37
Db	15625	CCAGCGCTGCTGGGGATCCATCGCCTGCGCCTCTCGGGCGGATATCCGAGAAACCCG	15566
QY	37	g-----Ly	38
Db	15565	CATCGAACAGGTGATGTCCCGCAGGAATCCACCTTCGCGCACGTAGTCTTGTCCCAACCG	15506
QY	38	silelleLeuAlaValMetAspArgGlnLysLysGluGluGluLysGlu-----GlnSe	56
Db	15505	CCTCCGGATCGGGGTCGTAATCGAAGGTGCATCCAGCGCGCAGGGAAGCCTCGATCG	15446
QY	56	rValLeuLyslleLysGluGlu-----	63
Db	15445	CATCTCTCCCTCTCCAGGAGACCCAGTAGTCTTCGGCGTAGCAGCGCCCCCCGCA	15386
QY	64	-----HisLys-----Al	66
Db	15385	GCCGCACGCCCATCATAGCATGCGATGGGCCCTGAACGCGAGATCCGATGGGTGGCCT	15326
QY	66	agLInProThrGlnTrpPheProPheSerGly-----IleThrGlu-----	79
Db	15325	GCAGCCTCTGCAACTCTTTTCAATGCACGGGTCAAGCCCTCAAGAAAGCGTAGGAGAGCGG	15266
QY	80	-----Leu-ValAsnAsnValLeuGlnProGlnGlnLysGlnP	92
Db	15265	CGCTCTATTATGAGAGATCTCTGAGTTCGAATCATTTGCAGAGCTTGCAA-----	15213
QY	92	roAsnGluLysGluProGlnThrLysLeuHisGlnGlnPheGluMet-----	107
Db	15212	-----GGCCAGCTGCTCGTCGCGGAGATCTTCAAAATCGGAATCGGAATCGGAACC	15167
QY	108	-----TyrLysGluGlnValLysLysMetGlyGluGluSerG	120
Db	15166	GGAACCTCGACGGATGTTTTCGACGGCCCCCGCAGCTCCAGAGCTTGGGAGAGGCCT	15107
QY	120	InclnGlnGlnGluGlnLysGlyAspAlaProThrCysGlylleCysHisLysThrLysP	140
Db	15106	GNACACGCACGAGACTACGGAGGCTGAGCAT-----GAACGCGCGCAATT	15059
QY	140	heAlaAspGlyCysGlyHisAsnCysSerrTyrCysGlnThrLysPheCysAlaArgCysG	160
Db	15058	TCCTCGTCGACATGAGTCCGCTGTACG-----	15030
QY	160	lyGlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysG	180
Db	15029	-----CAGCAGATTCGGGGCTCACCGTTTCGCAATT-----TGAC	14993

QY 483 uAspArgProAspMet-ArgArgAlaAspSerLeuArgLysGlnHisHisLeuAspProS 503
Db 14054 AGCTCGCGGCCCCCGCGGCTCCAGCCGCGCGAGCTGACCAAGGTGCGCGCACC 13995
QY 503 erSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsnAspSerL 523
Db 13994 CCGTGGCGCGCACCAGGTGCGCGCGAGCCGCTGACCCAGCTC----- 13951
QY 523 euSerSerAspGlnSerGluSerValA-GProProProArgProHisLysSerLysL 543
Db 13950 -----CCCTGTCCCGCGCTCACCAGCAC----- 13927
QY 543 ysGlyGlyLysMetArgGlnValSerLeuSerSerGluGluGluLeuAlaSerThrP 563
Db 13926 -----CGTCCCGCGAGGTCAGCTCGCGGCTCGGTGAGCTCTTC----- 13885
QY 563 roGluTyrThrSerCysAspAspValGluLeuGluSerValSerGluSerGluLysGlyA 583
Db 13884 --CGCTACCGCTGTGC-----GCGC 13866
QY 583 spSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSerAnThrA 603
Db 13865 ACCAGCGCGCGCGCAGCGAGCGCGCGCGCGCGCAAGCTC-----GCGCTCC 13815
QY 603 rgSerGluArgGlnLysLysArgMetTyrTyrGlyLysHisSerLeuGluGluAspLeuG 623
Db 13814 GCGCGCTGCGCAGCGCGCGCGCGCAGCAGCCAGCGTCCAGCGGCTCGGTCCCCATCC 13755
QY 623 luTrpSerGluPro----- 627
Db 13754 ATCAGCGCAGCGCGCTCGGGTCTCGCTTCGCGCGCGCGCGAGCAGCCCCCAGCAGC 13695
QY 628 -----GlnLeuLysAspSerGlyValAspThrCysSerSerThrThrLeuAeng 644
Db 13694 GCGCGCGCGCGCAGGTCTCGACGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCTG 13635
QY 644 luGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAspA 664
Db 13634 ATCCACACAGCTCGGTTCGAGCGCGCGCTCCGACAGCCAGCTTCGACGAGCGAC 13575
QY 664 rgLeuIleGlyArgIleLeuLeuAsnLys-----ArgLeuLysA 677
Db 13574 AGCGCTGCTGCTCGCTCATGCGAGCAGCCCGCGCCAGCTCCACCGGCTCGGCTCGG 13515
QY 677 spGlySerValPro-----ArgAspSerGlyAlaMetLeuGlyLeuLys----- 691
Db 13514 GCGGTGACGTGACACACACCGCTCAGCGCGCTCGCGCGCTGCTCGAGCGCGCACACC 13455
QY 692 -----ValValGlyLysMetThrGluSerGly--ArgLeuCysAlaPheIleTh 708
Db 13454 AATGATCGAGATCGGCAATCGCTCCGCGCGCGCGCGCTTCGCGCGCGCGCTCTCTCC 13395
QY 708 rlysValLysLys----- 712
Db 13394 TTCGTGACCATGACACACAGCGAGCCGCTCACCAGGAGGTGCTGCGAGGTCCACA 13335
QY 713 -----GlySerLeuAlaAspThr--ValGlyHisLeuArgProGlyAspGluValLeuG 730
Db 13334 GGTGGAAGCTCACCTGATACATATGTTGGCCACACAGCT-----GAACGCGGTGTC 13284
QY 730 uTrpAsnGlyArgLeuGlnGlyAlaThrPheGluGlu----- 743
Db 13283 CGCAGCTGCTCGCGCTCGCGCGCGCGCAGACGTAGAGCGCGCGATGCTGCCACGACCTGC 13224
QY 744 -ValTyrAniIleLeuGluSerLysProGluProGlnValGluLeuValValSerAr 763
Db 13223 CCGGTGAGCTGCGCAGCTAGCAGCAAGCCGTGGCTGCTGTGCGCTTTGTTCTGCAATC 13164
QY 763 gProIleGlyAspIleProArgIleProAspSer-----ThrHisAl 777
Db 13163 TCCACTCGATCCGAAGCTCGTCCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 13104
QY 777 aglnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSe 797

Db 13103 AACGGCAACACACGCGCTTGGCCCTTGACATCTGAGAGAAGCTGCGACCATCGTGTGG 13044
QY 797 rValThrSerProMetSerProGlyMetLeuArgAspValProGlnPhe-LeuSer----- 815
Db 13043 AGCGCGGCTCATCAGCGCG-----GATGACCCCATAGCTTCGCGCTGCGCGCTG 12996
QY 815 ----- 815
Db 12995 CCCCTCGGCTCCCGGCGCAACACACCGCGCGTAAAGCTTCGCGCTCGACGCGCAGAGC 12936
QY 815 ----- 815
Db 12935 TCCACGAGGCTCGGAACACACGAGACCTAGTCGAGCCGCGCTCACGAGCGCTCGTAA 12876
QY 816 -----GlyGlnLeuSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValT 834
Db 12875 AACCGGAGAGGTCCACTCTCTCGCGCGCGCGCAGCGCAGCTCGACAGCTCTGTCGAGC 12816
QY 834 hrIleLeuGlyAlaLysAspLeu-----ProSerA 844
Db 12815 TCGCGCGGCGCGCGCGGATCTCTCGCTCAACACGCGCGTGGGTGCTGCGCCACCGC 12756
QY 844 rgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuProAspArgSerA 864
Db 12755 GCGTCTCGCGCGCTGCTGCGAGCTGCTGACAGCTTACTCTCGCGCG----- 12706
QY 864 spLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys----- 880
Db 12705 -----GCCCGCGCGCTCGCGCTCCGACCGACAGCTGACGCGCGCGCGCTCC 12654
QY 881 -----TrpAsnGlnThrPheIleTyrSerProValHisArgA 893
Db 12653 TCGCGCGCAGCACAGCGCTCGCGCGAGAACACAGCTCCGACAGCTCGCGCTCCACCGCG 12594
QY 893 rgGluPheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluG 913
Db 12593 CGCGCGCGCGCAGCGCGCTCCAG-----CATCCCGCTCGCGCGAGAAC 12549
QY 913 luGluSerGluPheLeuGlyIleLeuIleGluLeuThrAlaLeuLeuAspAspG 933
Db 12548 AC-----CACCTTGCACCAACCTGATGATCCCGA 12519
QY 933 luProHisTrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeu----- 949
Db 12518 AGCCACGATGCTCGCGCAGCGACAGCTCTCTGAAACAGGTGCTCCCTCGCGCTCGCGC 12459
QY 949 ----- 949
Db 12458 AGCTTCGCGCGCGCGAGCAGCGAGTGGCGCGCGCTTCAGCCCGCGCGAGCCACG 12399
QY 950 -----ProArgProSerProTyrLeuPro-----ArgArgG 960
Db 12398 TCGCGCGCGCGCTCGCGCGCTCCAGCGAGTAGCGTCCGCTGGAAGCGGTAGTCTGCGC 12339
QY 960 lnLeuHisGlyLysSerProThrArgArgLeuGlnA-rgSerLysArgIleSerAspSerG 980
Db 12338 AGCTCCACGCGCGCGCGCGCATGCGCGA-----CAGCACCTGCGCGCGCAGTCC 12291
QY 980 luValSerAspTyrAspCysGluAspGlyValGlyValValSer-----AspTyrArgH 998
Db 12290 ACTCGTGCCTTCGACGTGACCGTCCGCGCGCGCTCCAGCGCGCTCACAAGCGCGCTTCTCTCT 12231
QY 998 isAsnGlyArgAspLeuGlnSerThrLeuSerValPro-----GluGlnValMets 1016
Db 12230 GCTCTCGCGCGAGCTCGCCACAACGTCGCTCGCTCTCTCTCGCGCAGGACCGCGCGC 12171
QY 1016 erSerAsnHisCysSerProSerGlySerProHis-ArgValAspValIleGlyArgThr 1035
Db 12170 CCAGCGCGCACAGCAGCGCATCCGGA---CGCACTCGA----- 12135
QY 1036 ArgSerTrpSerProSerAlaProProGlnArgAsnValGluGln---GlyHisArg 1054

12134	CGTATGTGCTCACCC-----CGCGCGCGCGAAGCGTGCGCATCCCGCTCCAGGAGCGCA	12081
Qy	GlyThrArgAlaThrGlyHisTyrAsnThr-----IleSerArgMetAspArgHisArg	1072
Db		
12080	CGGCCTCGCGCACTGCCTCACCCAGTACTCGGCCGACATCAGCGATCTTTCGCCACCGA	12021
Qy	ValMetAspAspHisTyrSerSer	1080
Db		
12020	GCTCGCGCGTCAAGCGCGCTCACCAAGGGTACTCGCGCTCCCATACGCGCACTCCCGCG	11961
Qy	-----AspArgAspArgAspCys-----	1086
Db		
11960	CCACCTCCCGTACTCTCTCCAGCATCCCGCTCCATCTGGCGGTGTGGAATGCTGCGACA	11901
Qy	-----GluAlaAlaAspArgGlnProTyrHisArgSerArgSerThrGluGlnArgPro	1104
Db		
11900	CGCGCAGACCGCGCTGCGCCCGCTCGCGCTCGAACCGCTGCCACCTCCACGACCG	11841
Qy	LeuLeuGluArgThrThrArgSerArgSerSerGluArg-----ProAspThrAsn	1122
Db		
11840	CGCCTT---CGTCCCGCTCAGCAGCGTCTGGCTGGACCGTGTATGCCCGATCCCA	11784
Qy	LeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAlaLeu	1142
Db		
11783	CGCGGCCTCCAGCCCATCG---CCGACAGCGCCCTCGACCTTCGGCTCCGACGGTT	11727
Qy	SerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrProGlyThr	1162
Db		
11726	-----CCACCGACATCATCGTCCGCGC-----CGCGCTGGCACCTCTGCATCA	11685
Qy	GlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThr-----	1177
Db		
11684	GCCAGCGCGGGCGCACATAGCTTCGCTCGCTCGCAAGGCTCAGCACGCCGCCACGT	11625
Qy	LeuGluArgSerAlaMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyr	1197
Db		
11624	GCGCAGCGCTCAGCTCTCT-ATCAGATGCCCGACG-----	11590
Qy	LysGlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSer	1217
Db		
11589	---AGGCGCGGGCGTCACTCCCAAGCCTCC-----CACTCGCGGTACAGCGCC	11542
Qy	GlyTyrAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspVal	1237
Db		
11541	ACTTCAGCGCAACACACCGGCTCGGTGTACTCGCTCCAGCAGCGGCTCTCTGC	11482
Qy	SerAspValSerAlaValSerArgThrSerSerAlaSerArg-----PheSerSer	1254
Db		
11481	TGCTGCCCGCTTCGCGCAACACCACTCTCTCAACCCACGGTCGAGATGCGGTCCAGC	11422
Qy	ThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPhe	1274
Db		
11421	GCTCGCACCTCTCGACGACCGCGACGAGACACGGGGTACACTTGTAAAGTCTCTTC	11362

RESULT 7

```

US-08-759-444-2/c
; Sequence 2, Application US/08759444
; Patent No. 5824309
; GENERAL INFORMATION:
; APPLICANT: DasSarma, Shiladitya
; APPLICANT: Morshed, Razeela
; APPLICANT: Stuart, Elizabeth
; APPLICANT: Black, Samuel
; TITLE OF INVENTION: RECOMBINANT GAS VESICLES AND USRS THEREOF
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/759,444
  FILING DATE: 05-DEC-1996
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 60/008,200
    FILING DATE: 05-DEC-1995
  ATTORNEY/AGENT INFORMATION:
    NAME: Meiklejohn, Ph.D., Anita L
    REGISTRATION NUMBER: 35,283
    REFERENCE/DOCKET NUMBER: 07880/003001
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 617-542-5070
    TELEFAX: 617-542-8906
  INFORMATION FOR SEQ ID NO: 2:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 8878 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: DNA
US-08-759-444-2

Alignment Scores:
Pred. No.: 3.36e-09
Score: 283.50
Percent Similarity: 33.31%
Best Local Similarity: 22.17%
Query Match: 3.42%
DB: 1
Length: 8878
Matches: 346
Conservative: 174
Mismatches: 570
Indels: 472
Gaps: 74

```

US-09-617-099B-1 (1-1590) x US-08-759-444-2 (1-8878)

Qy	10	ArgProAla	pro---ThrPro	AlaAla-----SerGlnPro-PropGlnProGlu	Me 26
Db	3797	AGAACTGCTCCAG	AGATCTCGCGGCTGTTGGGCAGACCTGAGTCGCGGTACAGTATA	3738	
Qy	26	tProAspLeu	SerHisLeu	ThrGluGluArgIysIleIleLeuAlaValMetAsp	46
Db	3737	CCGCGATTTAAAT	GACCTTCGACGTCGAA-----GGTGACTTGAGGT	3696	
Qy	46	gGlnIysLysGlu	GluLysGluGlnSerValLeuLysIleLysGluGluHis	ysAl 66	
Db	3695	ACAGAAATTCAG	CAACGCCAAA-----GTGTATCGCCTCTCCGATCTCGAAGGAC	3645	
Qy	66	a-----GlnPro	ThrGlnTrpPheProPheSerGlyIleThrGlu	Le 80	
Db	3644	GTTCAACCGAAT	AGACCATGTTGATCAGTTGCTGCTGTTTCGTCGTACTCAAGGC	3585	
Qy	80	uValAlan	AsnValLeuGlnProGlnGlnLysGln-Prob	AsnGlnLysGluProGlnThrL 100	
Db	3584	GGTATGACCGAT	TCAGACGCGTCATCACAAATCCCAAGAGAGAGACCAATGACTG	3525	
Qy	100	ysLeuHisGln	GlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerG	120	
Db	3524	AGAACTATACAT	ACAGTACGATATCATCGACAGAGAGATCTCGAATAGATGTCGAAGCGC	3465	
Qy	120	lnGlnGlnGln	GlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysP	140	
Db	3464	TTGCGGAGCGGA	ACAG-----GTCTATACGGTCGATCTACAGACACTCT	3420	
Qy	140	healaaSpGlyCys	GlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysG	160	
Db	3419	COGCT-----GTC	GCTCTCTGACATTTGATACGACCCGACCCCGAGCGCACCG	3375	
Qy	160	lyGlyArgValSer	LeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysG	180	
Db	3374	ATGAGGACGTGA	AGCTCATACACACGTGCTCCAGGAGGTATTGAAGCACGAGAGAAAC	3315	
Qy	180	lnGlnGluIleLeu	ThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuG	200	

```
Db 3314 GCACAGTCTCCCGATGAGCTTCGGGATGGCGTTC-:::|||||:::
Qy 200 InGlnProAspGlnLysValProArgGlyLeuArgAsnGlu-:::|||||:::
Db 3260 AGGGTGATTGCGCGGGCGGACGTGCTATTCGGCAGTACGCTGATGACATCGAGGAA 3201
Qy 216 roGlnGlnLysLysAlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspL 236
Db 3200 CGTCCGAACCTCGCGTGAAGATACTC-:::|||||:::
Qy 236 euSerValProAlaValGlu--LysGlyArgAlaHisGlyLeuThrArgGlnAspThrI1 255
Db 3161 ATACAGTCCCTCGAGAGAAATCCAGGAAACGTACGGATCACTAGCAGA-:::|||||:::
Qy 255 elySAsnGlySerGlyValLysHisGlnIleAlaSerAspMetProSerAspArgLysAr 275
Db 3109 -:::|||||:::
Qy 275 gSer-----ProSerValSerArgAsp-----282
Db 3083 ATCTCTTCACAGACCGCTGATCATCAATAAGTCGTACTCTCGTCACTTCGAGAAACGGC 3024
Qy 283 -----GlnAsnArgTyrGlnGlnSerGluGluArgGluAsp-----295
Db 3023 ATGCTTTTCACTCCGCTCATCGATGTCGAAGCGGAATACGACGAATCACTGATTCAGT 2964
Qy 296 -----TyrSerGlnTyrValProSerAspGlyThrMetPr 307
Db 2963 ACACGGGGCGTGGCGCGGTACATATTCGTGGACATTCACATCGCGCGGACAGCAAC 2904
Qy 307 oArgSerProSerAspTyrAla---AspArgArgSerGlnArgGluProGlnPhe-----324
Db 2903 AAGGAGCGCGATAATCATGTTTCATCATAGACGATCTCTGTGTAGCCGTTTCTCTCT 2844
Qy 325 -----TyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGlyHisAr 342
Db 2843 TGTGCTGATATCTCCAGAC-:::|||||:::
Qy 342 gHisSerLysGluTyrIleValAspAsp-GluAspVal---GluSerArgAspGluTyrG 361
Db 2801 ACACCAA-----ATCGATTTCGAGACACATCAAGGAGAACCACTCTCTGTATG 2754
Qy 361 lu---ArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuA 380
Db 2753 AAGTTGTTGAACGATCCGACGAGAGTACCAGCAGCGCAAGCA-:::|||||:::
Qy 380 laArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValS 400
Db 2702 AATCGCAACTGGAGATGGCGGAGCAGGTCCAGAAACAGATGCGTGACCGCATGGAGGTCA 2643
Qy 400 erArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuA 420
Db 2642 AGAAATAATGTTACCCGACGAA-:::|||||:::
Qy 420 spSerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluA 440
Db 2588 GCTCAGCGGCTGCTCGACCAATTCACACGCTCGTGGAGTCTCGCGACATCGAAGA 2529
Qy 440 rArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgTyrArgGluAlaGlnG 460
Db 2528 GGAGGGCGGTCAACGACACAGTCAAGCCGTATCGACCGGGGAAACGCCGAATTGATTA 2469
Qy 460 lyGlnSerSerTyrProGln-----ArgThrSerAsnH 471
Db 2468 CGACTACGAGTTTCCATCGGGGTGGACGAGCGGATCGTCTTACGACGAGGAACC 2409
Qy 471 isSerPro-----ProThrProArgArgSerProIleProLeuAspArgProAspMetA 489
Db 2408 ATCGTCCATTCGCTCCGATCCGACACAGCGATCGGTCAAGAGGACGA-----C 2358
Qy 489 rArgAlaAspSerLeuArgLysGlnHisLeuAspProSerSerAlaValArgLysT 509
```


Qy	357	gAspGluTyrGluArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspPr	377
Db	21917	AGGGGGGGTTCTCGCAGGCCATCGAGAGCGGGAATATCGACGCGTCTCAGCGCGCAGCTCC	21976
Qy	377	o-----AsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMe	393
Db	21977	ACGTGACGGCGACGACGAGCGCGCGCCCTTGC---CTGTCTCT---TCCACCCCTCG	22030
Qy	393	tArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSerLe	413
Db	22031	CGAGCTTTCGCCACGACGAGCGCAAGACGAGCGACGCTCGAGC-----CT	22075
Qy	413	uAla-----AsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspAr	430
Db	22076	GGGGCTACCGATCATCGTGAAGCTCTGACCACGC-----CACCAGCGCCGCG	22126
Qy	430	gProSerArgGlnArgSerValSerGluArgArgAlaMetGluAsnGlnArgSerTy	450
Db	22127	ACTTGGCGGCAC-----CTGGCTCTCTGTGCCCGCGCTCTGCAGCAGCAGCGCG	22180
Qy	450	rSerMetGluArgThrArgGluAlaGln---GlyGlnSerSerTyProGlnArgThrSe	469
Db	22181	TCGCTTCGCGCTCACCGAGCGCTCGCGCGCGCGCGCGCTCTCTCGCGTGC	22240
Qy	469	rAsnHisSerProProThrPro-----ArgArgSerProIlePro-----Le	483
Db	22241	TGAGCCAGGCGCCACCTGGAACCGGAGGCTCTCGCCGAGCACCTGCGCCAGGCTTCGCGCG	22300
Qy	483	uAspArgProAspMetArgArgAlaAspSerLeuArg-----495	
Db	22301	AGACCGCGCGCTCGCGGCGTCTCTGCTCTCTGCTCTGACGAAAGTCCCTCTCGCG	22360
Qy	496	-----LysG1	497
Db	22361	ACCATGCGCGGTGCGCGGACTGCTCTCTGCTCACCTCTCCAGCCCTCGCGCG	22420
Qy	497	nHisHis-----LeuAspProSerSerAlaValArgLysThr--	509
Db	22421	ACATCGCTCTGACGCGCCTTGTGGCTCTTACC CGCGCGCGCTCTCTCTCTCGGACACT	22480
Qy	509	-----509	
Db	22481	CCGACCCCATCGCCATCCGACGCGCGATGACCTGGGGCTGGCGCGCTGTCGCGCG	22540
Qy	510	-----LysArgGluLysMetGluThrMetLeuAr	519
Db	22541	TCGAGCACCCCGAGCGCTGGGAGGGCTGCTCGAGTGGCGCGCAGCATCGACGCGAGCG	22600
Qy	519	gAsnAspSerLeuSerSerAspGlnSerGluSerValArg-----Pr	533
Db	22601	CGTGGGCGCTGTCTCCCGTCTCGCCCTCGGACGATGAGGACGAGCTGCTCTCTCC	22660
Qy	533	oPro-----ProProArgProHisLysSerLysGlyGlyLysMe	547
Db	22661	GCCCGCGCGGTCTACGCTCGCGCTCTCTCGCGCTCCGCTCGG-----CGAGCG	22711
Qy	547	tArgGlnValSerLeuSerSerGluGluLeuAlaSerThrProGluTyrThr--	566
Db	22712	-CGCGCGCGCAGCTACCTTCAAGCCCGGAGCACCTCTCTCATCACGCGGACACCGCG	22770
Qy	567	-----SerCysAspAspValGluLeuGluSerGluSe	577
Db	22771	GCCTGGCGCTCAGTTCGCCGATGGCTGCTCGAGAGGGCGCAGACCTGTCTCTC	22830
Qy	577	rValSerGluLysGlyAspSerGlnLysArgLysThr-----591	
Db	22831	ATCAGCGCGAGGGGCCAGCGCGCGCTCGAGCTCCACGCCGAGCTCACGCGC	22890
Qy	592	-----SerGluGlnGlyValLeuSerAspSe	600
Db	22891	CTGGCGCGCGCTCATCTTCGCGCGTGTGATGTGCGCGCAGAGAGCGCTGTGCGCAGC	22950

[illegible]

Db 23278 GCGCGGCGGTATATTCACGGGGCCCCCTGGCAGCCAGCTGCGACCAAGCTGCTGTCG 23337
Qy 711 yLysGlySerLeuAlaAspThrValGly---HisLeuArgProGly-----AspG 727
Db 23338 CCGATGGCCCTTGGTGGCGGTGGCGGCTGCGCAAGCCCTGGACCAAGAGACC 23397
Qy 727 luValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnI 747
Db 23398 ACCGTACCGTCGCGACATGACTGGCGCGCTTGGCCCTTCGATCAGCTCGCTCGC 23457
Qy 747 le-----lleuGluSerLysProGluProGlnVal-----GluLeuValV 761
Db 23458 TCCCGCGCTCTCGCGCGACTTGGCGGAGCAGCGCGCTCGAAGACAGAGAGCGGCT 23517
Qy 761 alSerArgProIleGlyAspIleProArgIleProAspSerThrHisalaGlnLeuGluS 781
Db 23518 CTCTCTCGAGCAGCGCGCGCCCCGACCTCTCG-ACAAGCTCGGAGCGCTCGGA 23576
Qy 781 erSerSerSerPheGluSerGlnLys-----MetAspArgProSerIleSerV 798
Db 23577 GAGCGAGCAGCTCTGCTGCTCGCGCGTGTGTGCGAGAGAGCGGCTCTGCTCGG 23636
Qy 798 alThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnL 818
Db 23637 CCACGAAGCGCGCTTCCAG-----CTCGACCCGACAGAGCTTCTCGACCTCGG 23687
Qy 818 euSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyA 838
Db 23688 TCTCGATTCTGA-----TCATGACCGTCGAGCTTCTGCGCG 23723
Qy 838 laLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrP 858
Db 23724 CTTCACAGCAGCCACCGGCATCAAGCTCCCGGCCACCTCGCTCGACCATCCCTCTCC 23783
Qy 858 heLeuProAspArgSerAsp----- 864
Db 23784 TCATCGCTCGCGCTCTTCATGCGCGACTGCTCGCCCAAGCTCGGACAGGCTCTC 23843
Qy 865 -----LysAsnLysArgGhrLysThrValLysLysThrLeuLupProLysTrpAsnG 883
Db 23844 CGCGAGCGAGCGCGCGCGCTCGCGCGCGCTCGGAGCAGCAGCCCATCGCATCGT 23903
Qy 883 lnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluIleT 903
Db 23904 CGGCATGCG--CTCGCGCTCGCGCGCGCTCGCGGATGTCAGCTCTTTGGAGTTC 23961
Qy 903 hrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGluIleLeu- 922
Db 23962 CTCACCAAGGGCGCGAGCGGTGAGCCATTCCACAGAGCGCTGGGAGCGCGGTGCC 24021
Qy 923 -----lleGluLeuGluThrAlaLeuLeuAsp----- 931
Db 24022 CTCACGACCCCGACCGCGAGCGCGAGAGCTACGTCGCGCATCGCGCATGCTC 24081
Qy 932 -----AspGluProHisIleTyrLysLeuGlnThrHisAspValSerSerLeuProLeup 950
Db 24082 GACCAGATCGACTTTCGACCTTCGACCTTCTTCGCAATCGACCCCGCGGAGCGCAACAC 24141
Qy 950 roArg-----ProSerProTyrLeuProArgGlnLeuHisGlyGluSer----- 965
Db 24142 CTCGACCCCGACGACCGCTCTCTCGAATCTGCTGCTGGTGGCCCTCGAGAGCGCGGC 24201
Qy 966 -----ProThrArgArgLeuGlnArg-----SerLysA 975
Db 24202 ATCGTCCCGACCTCCCTCAAGGACTCTCCCTCACCGCGCTCTCTGTCGCGCATCGCGCGC 24261
Qy 975 rgIleSerAspSer-GluValSerAspTyrAspCysGlu----- 987
Db 24262 GAATACCGGATGCAAGAGCGGAGCTCGGAAGTTCCGAGGTTTACTTCATCCCAAGGCAC 24321
Qy 988 -----AspGlyValGly-ValValSerAspTyrArgHisAsnGlyAr 1001
Db 24322 TCCGCGTCTTTGGCGCGCGGGGCTTGGCTATACGCTCGGGCTCCAGG-----GGCGG 24375
Qy 1001 gaspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSe 1021
Db 24376 CGATCTTCGGTCGACACCGCGCTCTCTCTCGCTCGCTCTCTCTCTCTCTCTCTCTCT 24435
Qy 1021 rProSerGlySerProHisArgValAspValIleGlyArgThr---ArgSerTrpSerPr 1040
Db 24436 GGCCTCCGACAGGGGAGTGCACCTCGCCCTCGCGCGGGGGGTCTCGCTCTATGCTCTCC 24495
Qy 1040 oSerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrG1 1060
Db 24496 CCCCAGACCTTCGTCATCTTCGCTTCGCGCTTCGCGCGCTGGCGCGCGCGCTCCAG 24555
Qy 1060 yHisTyrAsnThrIle-SerArgMetAspArgHisArgValMetAspAspHisTyrSers 1080
Db 24556 ACCTTCTCGGACACCGCGAGCTTACGAGCGGAGAG----- 24593
Qy 1080 erAspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSert 1100
Db 24594 --AGCGTCTGCTCTTCCTTCCTTCGAGCGGATCGG----- 24626
Qy 1100 hrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProA 1120
Db 24627 -----CGAGCGCTTCGCGCGAGAC-ACCGCTCTCTCTCTCTCTCTCTCTCTCTCT 24680
Qy 1120 spThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerP 1140
Db 24681 CAACACGACGCGCGCT-----CGAGCGTATCACGCGCC 24716
Qy 1140 roAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrP 1160
Db 24717 CAACGCGACCTCCGAGAGAGTCTCTCGCGCGCGCTCCACGACGCGCATCA----- 24772
Qy 1160 roGlyThrGlyArgGlyArgGlnLeuProGlnLeuProGlyGlyThrLeuGluA 1180
Db 24773 -----CCCCCGCGAGCTCGAGCTGCT 24794
Qy 1180 rgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnV 1200
Db 24795 CGAGTGCCATG----- 24805
Qy 1200 alAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpA 1220
Db 24806 -----GCACGCGCACCTCTCTGGGAGACCCCATCGAGTGCAG-----CCCTGGC 24851
Qy 1220 spProHisArg---GlyAlaAspThrValSerThrLysSerSerAspSerAspValSerA 1239
Db 24852 CGCGTCTACGCGACGCGAGACCGCTGAAAGCTCTCTCTCTCGCGCGCTCAAGAC 24911
Qy 1239 spValSerAlaValSerArgThrSerSerAlaSerArgPhe----- 1252
Db 24912 CAACATCGGCGCATCTCGAGCGCGCTCGCGCTCGCGCGCTCGCAAGATGGTTCGCTC 24971
Qy 1253 --SerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleS 1272
Db 24972 GCTCGCGACGCGCGCTCGCGCGCGCTCGCGCGCGCTCGCGCGCGCTCGCGCGCGCT 25016
Qy 1272 erValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrL 1292
Db 25017 CAATCCCTCATCG-----AGTGGAGCGCTCGCCAT 25049
Qy 1292 yssSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerA 1312
Db 25050 CGAGCTGCTGATACCCCGAGGCTT---GGCCCC-CGCGACGAGATGGCGGCTCCCGCGC 25105
Qy 1312 spThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgArgSerSerIleGlyAlaL 1332
Db 25106 GCGCGCGCATCTCGCGCTCTCGGATTCGCGGACCAAGCGCGCGCTCATCTCGAGAGG 25165
Qy 1332 yMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeu----- 1349
Db 25166 CTCCCGCGCGCTCTGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCCCG 25225

Qy	1350	-----SerGlnThrGluGlyGlyGlyLysL	1358
Db	25226	CGCGTGGCCGCTCCTGTCCGCCAGGAGCGCCGCGCTCGCGCCAGCGCAAGC	25285
Qy	1358	ysLeuArgSerThrValGlnArgSerThrGlyLeuValGluMetArgant	1378
Db	25286	GGTCCGCGACCACTCGTCCGCCACGACGACTCACCTCCGCGATGTGCCTATTGCG	25345
Qy	1378	rpMetThrArgGlnAla-----SerArgGluS	1387
Db	25346	TGGCCACCAACCCGCGCCCACTTCAGCACGCGCGCTCTCGTAGCCACCAACCGCGAAG	25405
Qy	1387	erThrAspGlySerMetAnSerTyrSer-----	1396
Db	25406	AGTCTCTCTCCGCGCTCGACTCGCTCGCCACGACGACGCGCCGCCGAGCACCGTCTCTCG	25465
Qy	1397	-----SerGluGlyAsnLeu-----IlePheProGlyVal-----	1406
Db	25466	GACGAGCGGAAGCACGCGCAAGCTCGTCTTCCTTCCTGGGCAAGGCTCGACGGG	25525
Qy	1406	-----	1406
Db	25526	AAGGGATGGCCCTCTCGCTGCTCGACTCTCGCCGCTCTTCGCGCTCAGCTCGAAGCAT	25585
Qy	1407	-----ArgLeuAlaSerAspSerGlnPheSer-----	1415
Db	25586	GCGAGCGCGCTCGCTCTCTCAGTCGATGAGGAGCTGTGCGCTCTGCGCCGCGAAG	25645
Qy	1416	-----AspPheLeuAsp-----	1419
Db	25646	AGGGCGCCCTCCCTCGACCGGCTCGACGTCGTACAGCCGCCCTCTTTCGCGTCATGG	25705
Qy	1420	-----GlyLeuGlyProAlaGlnLeuValGlyArgG	1430
Db	25706	TCTCCTCGCGGCCCTCTGCGCGCTCGCTCGGGGTAGAGCCGCCGCTCGTGGCCAC-	25764
Qy	1430	lnThrLeuAlaThrProAlaMetGlyAspIleGln-----ValGlyMetMetA	1446
Db	25765	-----AGTCAGGGCGAGATCGCGCGCCCTTCGTCGAGCGCGCTCTCT	25807
Qy	1446	spLysLysGlyGlnLeuGluValGluIleArgAlaArgGlyLeuValLysProG	1466
Db	25808	CCCTCGAGGAGCGGCCCGCATCGCCCTGGGAGCAAGCGCTCACCCGTCGCG	25867
Qy	1466	lySerLysThr-----LeuProAlaProTyrValLysValTyrLeuLeuA	1481
Db	25868	GCAACGGGCGCATGCGCGCGCTGAGCTCGGCGCTCGACCTCCAGACCTACCTCGCTC	25927
Qy	1481	spAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeu-----	1498
Db	25928	CCTGGGGGACAGGCTCTCATCGCGCGCTCAACAGCCCCAGGCCACGCTCGTGTCCG	25987
Qy	1499	--GluPro-----LeuTyrGlnGlnLeuLeuSerPheGluGluSerProG	1513
Db	25988	GCGAGCCCGCGCCATCGACGCGCTGATCGACTCGCTCACGCGCGAG-----	26037
Qy	1513	lnGlyArgValLeuGlnIleValTrpGlyAspTyrGlyArgMetAspHisLysSerP	1533
Db	26038	-----GTCTCGCCCGAAAGTCGCGCTCGACTACGCCCTCCC	26074
Qy	1533	heMetGlyValAlaGlnIleLeuLeuAspGluLeuGlu-----LeuSerAnMet-Val	1550
Db	26075	ACTCCGCCAGATGGAGCGGCTCCAGACAGCTCGCGAGGTCTACCAACATCGCTC	26134
Qy	1551	IleGlyTrpPheLysLeuPheProProSerSerLeu---ValAspProThrSerAlaPro	1569
Db	26135	CTCGGA-----CGTGGAGCTCCCTCTTATTTCGACCGCTACCCGCGACCA	26179
Qy	1570	LeuThrArgArgAlaSer-----GlnSerSerLeuGluSerSerThrGlyProSer	1586
Db	26180	GGCTCCAGCGGCTCCGAGCTCGACGGCGCTATGTGTATCGAACTCCGGCAACCCGTCC	26239

; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10532
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/457,646A
 ; FILING DATE: 01-JUN-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/457,205
 ; FILING DATE: 01-JUN-1995
 ; APPLICATION NUMBER: 08/258,261
 ; FILING DATE: 08-Jun-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Elmer, James Scott
 ; REGISTRATION NUMBER: 36,129
 ; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 919-541-8614
 ; TELEFAX: 919-541-8689
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 28958 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; US-08-457-646A-6

Alignment Scores:
 Pred. No.: 1.36e-07 Length: 28958
 Score: 270.50 Matches: 436
 Percent Similarity: 30.17% Conservative: 194
 Best Local Similarity: 20.88% Mismatches: 768
 Query Match: 3.26% Indels: 696
 DB: 1 Gaps: 90

US-09-617-099B-1 (1-1590) x US-08-457-646A-6 (1-28958)

QY 132 CysGlyIleCysHisLysThrLysPheAlaAspGlyCysGlyHisLysCysSerTyrCys 151
 Db 20882 TGTCTGATACCTCGCGCGTCTTCC-----GGGCACAGCTCGAAGCGTGG 20926
 QY 152 Gln-----ThrLysPheCysAlaArgCysGlyGly-----Arg 162
 Db 20927 AGCGCGCCCTCGGCGCCACAGTGGTCTCTCGCGGTCTCCCGCGAGAGG 20986
 QY 163 ValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGlnGlu 182
 Db 20987 GCGCGCCCGCTCGACCGGTCGAGTGTCCAGCCCGCGCTGTCTCGATGATGCT 21046
 QY 183 IleLeuThrLysSerGlyAla-----189
 Db 21047 CGCTGGCCGCTGTGGCGCTCCATGGCGTCGAGCCGCGGTGTGGCCATAGCC 21106
 QY 190-----TrpPheTyrAsnSerGlySerAsnThrLeu-GlnG 201
 Db 21107 AGGCGAGATCGCGCGGCTGTGTGGCGGCGCTGTCTCGAGGACGTCGCCAAGC 21166
 QY 201 nProAspGlnLysValPro-----ArgGlyLeuArgAsnGluAlaProGlnGlu 219
 Db 21167 TGGTGGCGCTCGCGACCGCTGCGCTGTGAGCTCGCCGCCAGGGGCGC-----21215
 QY 219 sLysAlaLysLeuHisGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValPr 239
 Db 21216-----CATGGCGCGGTGGAGCTGCGGAGGCGAGGTGCGACGCGGCTCC 21262
 QY 239 oAlaValGluLysGlyArgAlaHisGlyLeuThrArgGln-----AspThr1 255
 Db 21263 AGCGCTATGGCGATCGGCTCTCCATCGGGCGATCAACAGCCCTCGTTTCACGACGATCT 21322
 QY 255 eLysAsnGly-----SerGlyValLysHisGlnIle-- 265
 Db 21323 CGGCGAGCCCTCGCGTCGCGCCCTGCTCGCGATCTGAGTCCGAGGCGCTTCG 21382
 QY 266-----AlaSerAspMetProSerAspAr 273
 Db 21383 CCTCAAGCTGAGTTACGACTTCGCTCCACTCCGCGCAGGTGCGAGTCGATTCGCGACG 21442
 QY 273 g---LysArgSerProSerValSerArgAsp-----G 283
 Db 21443 AGCTCCTCGATCTCTGTCTGTGCTCGCGCGCTCGACGCGCTCGACGCGTCTTACTCCA 21502
 QY 283 nAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrValProSerAs 303
 Db 21503 CGGTGAGCGCGCGCGGATCGACGCGGAGCGAGCTCGACCGCGCTACTGTACCGAACC 21562
 QY 303 p-----303
 Db 21563 TCGGCGAGCGGTCCGCTTCGCGACGCTGTGCAAGCGCTCTTTCGCGGAGAACATCGCT 21622
 QY 304-----GlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 321
 Db 21623 TCTTCGTGAGGTGAGCCCGAGTCTGTCTGACCTTGGCCTTGCACGAGCTCTCGAAG 21682
 QY 321 uProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgGlyHis 341
 Db 21683 CGTCGAGCGCTCGGCGGCGGTGTGCGTCTCTCTGTG-----GAGCGAGAGGGGATC 21736
 QY 341 s-----ArgHisSerLysGluTyrIleValAs 350
 Db 21737 TAGCGCGCTTCTCGTCTCGCTCTCGAGCTCTAGTCAACGCGTTCGCGCTTGGATGGA 21796
 QY 350 pAspGluAspVal-----354
 Db 21797 CGACGATCTGCGCCCGCGGAGCGGTGCGCTGCCACCTACCCCTTCAGCGCGAGC 21856
 QY 355-----GluSerAr 357
 Db 21857 GCTTCGGCTCGACGCGCTCCAGGACCGCGCGCGGTGCGTCCACCTTCAGCGCGAGC 21916
 QY 357 gAspGluTyrGluArgGlnArgGluGluTyrGlnAlaArgGluTyrArgSerAspPr 377

Db 21917 AGGGCGGTTCTGGCAGGCCATCGAGCGGGAATATCGACGCGCTCGACGCGCCAGCTCC 21976
Qy 377 o-----AenLeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMe 393
Db 21977 ACGTGACGCGCAGCAGCGCGCGCGCTTGC---CTGCTCT---TCCGACCCCTCG 22030
Qy 393 tArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSerLe 413
Db 22031 CGAGCTTTCCGCACAGCGCGCAAGAGCAGGCGACGGTCGACGC-----CT 22075
Qy 413 uala-----AenAlaGluLeuGluAspSerArgIleSerLeuArgMetAspAr 430
Db 22076 GGCCTTACCGCATCAGTCGGAAGCTCTGACACCGC-----CACCAAGCGCGCGC 22126
Qy 430 gProSerArgGlnArgSerValSerGluArgAlaAlaMetGluAenGlnArgSerTy 450
Db 22127 ACCTGGCGGCAC-----CTGGCTCTCTGCTGTCGCGCGCTCTGACACGACGCGC 22180
Qy 450 rSerMetGluArgThrArgGluAlaGln---GlyGlnSerSerTyrProGlnArgThrSe 469
Db 22181 TCCCTTCGCGCTCACGAGCGCTCGCCGCGCGCGCGCGCTCTCTGCGCGTGC 22240
Qy 469 rAsnHisSerProProThrPro-----ArgArgSerProIlePro-----Le 483
Db 22241 TGAGCCAGGCCACCTGACCGCGAGGCTCTCGCGAGCACTCGCGCAGGCTTGC 22300
Qy 483 uAspArgProAspMetArgAlaAspSerLeuArg-----495
Db 22301 AGACCGCGCGCTCGCGCGCTCTCTGCTCTCTGCTCTCGACGAAAGTCCCTCGCGC 22360
Qy 496 -----LysG1 497
Db 22361 ACCATGCGCGCTCGCGCGGACTCGCTTCTCGCTCACCTCGTCCAGCCCTCGCGC 22420
Qy 497 nHisHis-----LeuAspProSerSerAlaValArgLysThr-- 509
Db 22421 ACATCGCCTCGACGCGCCTTGTGGCTCTTACCGCGCGCGCTCTCGTTCGACACT 22480
Qy 509 -----509
Db 22481 CGGACCCCATGCGCCATCGCAGCGAGCGATGACTGGGGCTGGCGCGCTCGTTCGCGC 22540
Qy 510 -----LysArgGluLysMetGluThrMetLeuAr 519
Db 22541 TCGAGCACCCCGAGCGCTGGGAGGGCTCGTCGAGCTCGCGCGCAGCATCGACGCGAGC 22600
Qy 519 gAsnAspSerLeuSerSerAspGlnSerGluSerValArg-----Pr 533
Db 22601 CCGTGGCGCGCTTGTCTCCGCTCTCGCCCTCGCGCAACGATGAGGACCGACTCGCTCTCC 22660
Qy 533 oPro-----ProProArgProHisLysSerLysLysGlyLysMe 547
Db 22661 GCGCGCGCGGTTTACGCTCGCGCGCTCGTCGCGCTCGCGCTCGCG-----CGACG 22711
Qy 547 tArgGlnValSerLeuSerSerGluGluGluLeuAlaSerThrProGluTyrThr-- 566
Db 22712 -CGCGCGCGCAGCTTACAGCCCGGAGGACCTCTCTATCATCCGAGGACCGCGC 22770
Qy 567 -----SerCysAspAspValGluLeuGluSerGluSe 577
Db 22771 GCGCGTGGCGCTCACGTCGCGCGATGCTCGCTCGAAGCGCGCAGACACCTCGTCTC 22830
Qy 577 rValSerGluLysGlyAspSerGlnLysGlyLysArgLysThr-----591
Db 22831 ATCAGCGCGGAGGCGCCAGCGCGGCGCTCGGAGCTCCACGCGCGAGCTCACGCGC 22890
Qy 592 -----SerGluGlnGlyValLeuSerAspSe 600
Db 22891 CTGGCGCGCGCTCACCTTCGCGCGCTGTGTGTGTCGCCGACAGGCGCTGTGCGCACG 22950
Qy 600 rAsnThrArgSerGluArgGlnLysLysArg-----MetTyrTy 613

22951 CTTCTCGAGCAGCTCGACGCGCAGAGGCTCGCAGTCCGCGCGCTGTTCACGCGCGCGC 23010
Qy 613 xGlyGlyHisSerLeuGlu-----AspLeuGluTrpSerGluProGlnLys-- 630
Db 23011 ATCGCGCGCCACGCTCGCTCGCGCGCACCTCTCTCATGAGAGTCGCGAGCTTGTCTCT 23070
Qy 631 -----AspSerGlyValAspThrCysSer-SerThrThrLeuAenGluG 645
Db 23071 GCCAAGGTCTTAGCGCGCAGGAACTCCACGACCTGTCTCGTCTCTGACCCCTCGACGC- 23129
Qy 645 luHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAspArgL 665
Db 23130 -----CTTCTGCTCTTCTCGTCCATCGACGCGCTGGCGCGCGCA 23172
Qy 665 euIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSerG 685
Db 23173 CAACAGCCGGAATACGCGCGCGGAAACGCTT-----CCTCGACGCCCTG 23217
Qy 685 lyAlaMetLeuGlyLeuLysValValGlyLysMetThrGluSerGlyArgLeu--- 703
Db 23218 GCCGACCGCGCGCAGTCTTGGACACCGCAGCTCGTGTGTGTGGCGCGCTGGGCG 23277
Qy 704 -----CysAlaPheIleThrLysValL 711
Db 23278 GCGCGCGTGTATATTTCACGCGCGCCCTCGCAGCCAGCTCGAGCAACGCTGCTCTCG 23337
Qy 711 yLysGlySerLeuAlaAspThrValGly---HisLeuArgProGly-----AspG 727
Db 23338 CCGATGCCCTTCTGCGCGCGCTGGCGCGCTCGCGCAAGCTTGGACGACGACGAGACC 23397
Qy 727 luValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnI 747
Db 23398 ACCGTACCGCTCGCGACATCGACTGGCGCGCTTTCGCGCTTCGATCAGCGTCTGCTCG 23457
Qy 747 le-----IleLeuGluSerLysProGluProGlnVal-----GluLeuValV 761
Db 23458 TCCCGCGCTCTCGCGGACTTCCGCGCAGCGCGCGCTCGAAGACAGAGAGCGCGCT 23517
Qy 761 alSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGluS 781
Db 23518 CTTCTTCGAGCAGCGCGCGCGCGCTCTCTCG-ACAAGCTTCGAGCGCGCTCGGA 23576
Qy 781 exSerSerSerPheGluSerGlnLys-----MetAspArgProSerLyserv 798
Db 23577 GAGCAGCAGCTCTGCTCTGCTCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 23636
Qy 798 alThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnL 818
Db 23637 CCAGAGGCGCTTCCAG-----CTCGACCGCGACAGAGGCTTCTTCACCTCGG 23687
Qy 818 euSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrLeuGlyA 838
Db 23688 TCTCGATTGCA-----TCATGACCGCTCGAGTCTGTGTGTGTGTGTGTGTGTGTGTGT 23723
Qy 838 laLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrP 858
Db 23724 CTTGCAACAGCGCACCGGCATCAAGTCCCGCGCACCTCTGCGCTTCGACCATCCCTCTCC 23783
Qy 858 heLeuProAspArgSerAsp-----864
Db 23784 TCATCGCGTCCGCTCTTCATGCGCGACTCGCTCGCGCCACGCGCTCGGACGAGGCTCTC 23843
Qy 865 -----LysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLysTrpAsnG 883
Db 23844 CGCGAGGCGACCGCGCGCTCGCGCGCTCGAGCGACGAGCCCATCCCATCGT 23903
Qy 883 lnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluIle 903
Db 23904 CGGCATGGC--CCTGCGCTCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 23961
Qy 903 hrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGluIleLeu- 922
Db 23962 CTCACCAAGGCGCGCGCTGT 24021

Db 25928 CTGGGGGACAGGCTCTCCATCCGCGCGCTCAACAGCCCGCCAGGCGCCAGCTCGTGTCCG 25987
Qy 1499 --GluPro-----LeuTyrGlnGlnLeuLeuSerPheGluGluSerProG 1513
Db 25988 GCGAGCCCGCCGACGAGCGCTCATCGACTCGCTACCGCAGCGGAG-----26037
Qy 1513 InGlyArgValLeuGlnIleValTyrGlyAspTyrGlyArgMetAspHisLysSerP 1533
Db 26038 -----GTCTCGCCGAAAGTCCGCGTCTGACTACCTCTCC 26074
Qy 1533 heMetGlyValAlaGlnIleLeuLeuAspGluLeu-----LeuSerAsnMet-Val 1550
Db 26075 ACTCGCCCGAGATGAGCGCGCTCCAGAGAGAGCTCGCGCAGGTCTAGCCAACTCGCTC 26134
Qy 1551 IleGlyTyrPheLysLeuProProSerSerLeu---ValAspProThrSerAlaPro 1569
Db 26135 CTCGGA-----CGTGGAGCTCTCTTTATTCGACCGCTGACCGGACCA 26179
Qy 1570 LeuThrArgAlaSer-----GlnSerSerLeuGluSerSerThrGlyProSer 1586
Db 26180 GGCTCGACGGCTCCGAGCTCGAGCGCGCTACTGTGTATCGAACTCCGGCAAACGCTCC 26239
Qy 1587 TyrSerArgSer 1590
Db 26240 TGTCTCGAGCG 26251

RESULT 12

US-08-458-076A-6
; Sequence 6, Application US/08458076A
; Patent No. 5698425
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,076A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CSC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-458-076A-6
Alignment Scores:
Pred. No.: 1.36e-07 Length: 28958
Score: 270.50 Matches: 436
Percent Similarity: 30.1% Conservative: 194
Best Local Similarity: 20.8% Mismatches: 768
Query Match: 3.26% Indels: 696
DB: 1 Gaps: 90
US-09-617-099B-1 (1-1590) x US-08-458-076A-6 (1-28958)
Qy 3 AlaProLeuGlyProArg-----GlyArgProAlaProThrProAlaAlaSerGlnPro 20
Db 20577 GCTCTGTGCGCCAGAGCGAGCGCGCTGCGCGCCAGGCGCCAGAGCTCCGCGACCA 20636
Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIleLe 40
Db 20637 CCTCTCTCCGCCAGAGACCT-----CGCCCTGCGCGCGATGTAGCCTA 20678
Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysIle 60
Db 20679 CTGCGCT-----20684
Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 20685 ---CGCCACCAACCGCGGCTA-CCTCGAGCACCGCGCTCTCTGTGTCTCCAGACCGCG 20740
Qy 81 ValAsn-----AsnValLeuGlnProGlnGlnLysGlnPro-----92
Db 20741 AAGAGCTCTCTCCGCGCTCGATTGCTGCGCGCGAGGAGCGCGCCCGAGCACCGCTCG 20800
Qy 93 ---AsnGluLysGluProGlnThrLysLeuHisGlnGlnPheGluMetTyrLysGluGln 111
Db 20801 TCGAACGAGCGGAGCACCGCGAGCTGCTGCTGCTCTTC-----20842
Qy 112 ValLysLysMetGlyGluGluSerGlnGlnGlnGlnGlnLysGlyAspAlaProThr 131
Db 20843 -----CTGGGC-----AAGGCTCGCAGTGGAGGATGCGCTCTCC 20881
Qy 132 CysGlyIleCysHisLysThrLysPheAlaAspGlyCysGlyHisAsnCysSerTyrCys 151
Db 20882 TGCTCGATACCTCCGCGCTCTTC-----GGGCACAGCTCGAAGCGTCCG 20926
Qy 152 Gln-----ThrLysPheCysAlaArgCysGlyGly-----Arg 162
Db 20927 AGCGCGCCTCGCGCCCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20986
Qy 163 ValSerLeuArgSerAsnLysValMetTyrValCysAsnLeuCysArgLysGlnGlu 182
Db 20987 GCGCGCCCGCTCGACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21046
Qy 183 IleLeuThrLysSerGlyAla-----189
Db 21047 CGCTGCGCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21106
Qy 190 -----TyrPheTyrAsnSerGlySerAsnThrLeu-GlnG 201
Db 21107 AGGCGAGATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21166
Qy 201 nProAspGlnLysValPro-----ArgGlyLeuArgAsnGluAlaProGlnGlu 219
Db 21167 TGTGGCGCTCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21215
Qy 219 LysAlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValPr 239

Db 21216 -----CATGGCGCGTGGAGCTGCGAGGCGGAGGTGCGACGGGCTCC 21262
QY 239 oAlaValGluLeuGlyArgAlaHisGlyLeuThrArgGln-----AspThr11 255
Db 21263 AGCGCTATGCGGATCGCTCTCCATCGGGCGATCAACGCCCTCGTTTCACGACGATCT 21322
QY 255 eLysAendly-----SerGlyValLysHisGlnIle-- 265
Db 21323 CGGCGAGCCCTTGGCGTGGCGGCTGTCTCGCGATCTGGAGTTCGAGGGGCTCTCG 21382
QY 266 -----AlaSerAspMetProSerAspAr 273
Db 21383 CCCTCAAGCTGAGTACGACTTCCCTCCACTCCGCGCAGGTGCGATTCGCGACG 21442
QY 273 g----LysArgSerProSerValSerArgAsp-----GI 283
Db 21443 AGCTCTCGATCTCTCTGCTGCTCGAGCGCGCTCGAGCGGGTCCCTTCTACTCCA 21502
QY 283 nAsnArgArgTyrrGluGlnSerGluGluArgGluAspTyrrSerGlnTyrrValProSerAs 303
Db 21503 CGGTAGCGGCGCGCGATCGACGGGAGCGAGTTCGACGCGCTACTGTGTACCGGAACC 21562
QY 303 p----- 303
Db 21563 TCCGGCAGCGGTCGCTTCGACAGCGCTGTGCAAGGCTCTTTCGCGGAGAACATCGCT 21622
QY 304 -----GlyThrMetProArgSerProSerAspTyrrAlaAspArgArgSerGlnArgG1 321
Db 21623 TCTTCGTGGAGGTGAGCCCGACTCTGTGTGACTTGGCTTGGCAGAGCTCTCGAAG 21682
QY 321 uProGlnPheTyrrGluGlnProGlyHisLeuAsnTyrrArgAspSerAsnArgArgGlyHi 341
Db 21683 CGTCGGAGCGTCCGCGCGGTGTGCTCTGTGTG-----GAGCGACGAAGGGGATC 21736
QY 341 s-----ArgHisSerLysGluTyrrIleValas 350
Db 21737 TAGCGCGCTTCTCGTCTCGCTCTCCGAGCTCTACGTCAACGGCTTCGCCCTGGATTGGA 21796
QY 350 pAspGluAspVal----- 354
Db 21797 CGACGATCTTCCGCGCGGAGGCGGTGCGGTGCCACCTACCCCTTCACGCGCGAGC 21856
QY 355 -----GluserAr 357
Db 21857 GCTTTCGGCTGCGAGCTCCACGCGACCCGCGCGCGCGTCAACCACTTGTCTCGGCTCG 21916
QY 357 gAspGluTyrrGluArgGlnArgArgGluGluGluTyrrGlnAlaArgTyrrArgSerAspPr 377
Db 21917 AGGGCGGTTCGCGAGGCTATCGAGAGCGGATATCGACGCGCTCAGCGGCGAGCTCC 21976
QY 377 o-----AsnLeuAlaArgTyrrProValLysProGlnProTyrrGluGluGlnMe 393
Db 21977 ACGTGGACGGCGAGCAGCGCGCGCGCTTGC-----CTGCTCTCT--TCCACCCCTCG 22030
QY 393 tArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSerIe 413
Db 22031 CGAGCTTTCGCGAGCGCGCAGAGCGGCGCGCTGCGACG-----CT 22075
QY 413 uAla-----AsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspAr 430
Db 22076 GCGCTACCGCATCACGTGAAGCTCTGACACCGC-----CACACGCGCGCGC 22126
QY 430 gProSerArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTy 450
Db 22127 ACCTGGCGGCGAC-----CTGGCTCTCTGTGTGTGCGCGCGCTCTGACGACGACGCGC 22180
QY 450 rSerMetGluArgThrArgGluAlaGln-----GlyGlnSerSerTyrrProGlnArgThrse 469
Db 22181 TCCCTCCGCGCTCAGCGAGCGCTCCGCGCGCGCGCGGTCTCGCGTCTCGCGTGGCGC 22240
QY 469 rAsnHisSerProProThrPro-----ArgArgSerProIlePro-----Le 483

22241 TGAGCCAGGCCACCTGGACCGCGAGGCTCTCGCGAGCACCTGCGCAGGCTTGGCGCG 22300
QY 483 uAspArgProAspMetArgArgAlaAspSerLeuArg----- 495
Db 22301 AGACCGCGCGCTCTCGCGCGTCTCTCGCTCTCTCGCTCTCGCTCTCGCTCTCGCTCTCG 22360
QY 496 -----LysGI 497
Db 22361 ACCATGCCCGCTGCGCGCGGAGTCTGCTTCTGCTTCACTTCTGCTTCAAGCCCTCGGCG 22420
QY 497 nHisHis-----LeuAspProSerSerAlaValArgLysThr-- 509
Db 22421 ACATGCCCTCGACGCGCTTGTGTGCTCTTCAACCGCGCGCGCTCTCGCTCGGACACT 22480
QY 509 ----- 509
Db 22481 CGACCCCATCGCCCATCGACGAGCGAGTCACTTGGGCGCTTGGCGCGCTGCTCGCGC 22540
QY 510 -----LysArgGluLysMetGluThrMetLeuAr 519
Db 22541 TCGACACCCCGAGCGCTGCGGAGGCTCTGCTGCTCGCGCAGCGATCGACGCGCG 22600
QY 519 gAsnAspSerLeuSerSerAspGlnSerGluSerValArg-----Pr 533
Db 22601 CGGTGGCGCTTGTCTCCCGCTCTCGCTCGCAACGATGAGGACAGCTGCTCTCC 22660
QY 533 oPro-----ProProArgProHisLysSerLysLysGlyGlyLysMe 547
Db 22661 GCCCGCGCGGTCTACGCTCGCGCTCTGCTCGCGCTCTCGCTCGCTCGCTCGCTCG 22711
QY 547 tArgGlnValSerLeuSerSerGluGluGluLeuAlaSerThrProGluTyrrThr-- 566
Db 22712 -CGCCCGCGAGTACTTCAAGCCCGAGGACCCCTCTCATCACCAGGACACCGGC 22770
QY 567 -----SerCysAspAspValGluLeuGluSerGluSe 577
Db 22771 GCGCTGGCGCTCACGCTCGCCGATGCTGCTCGAGAGGCGCAGACGACCTCGTCTCC 22830
QY 577 rValSerGluLysGlyAspSerGlnLysGlyLysArgLysThr----- 591
Db 22831 ATCAGCGCGCGAGGCGCGAGCGCGCTCGAGCTCCACGCCGAGCTCACGGCGC 22890
QY 592 -----SerGluGlnGlyValLeuSerAspse 600
Db 22891 CTGGCGCGCGCTCACCTTCCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 22950
QY 600 rAsnThrArgSerGluArgGlnLysLysArg-----Metyty 613
Db 22951 CTTCGAGCAGCTCGACCGCGAGGTCGCGAGTCCGCGCGCTTCCACGCGCGCGC 23010
QY 613 rGlyGlyHisSerLeuGluGlu-----AspLeuGluTrpSerGluProGlnLys-- 630
Db 23011 ATCGGCGCGCAGCTCGCTCGCGCGCAGCTCTCTCATGAGCTCGCGACGTTGTCTCT 23070
QY 631 -----AspSerGlyValAspThrCysSer-SerThrThrLeuAsnGluG 645
Db 23071 GCCAAGTCTAGCGCGCAGGAACTCCAGACCTGCTGCTCTGCTCTGACCCCTCGACGC- 23129
QY 645 luHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAspArgL 665
Db 23130 -----CTTCGCTCTTCTTCTGCTCATCGACGCGCTTGGGCGCGCGA 23172
QY 665 eulleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSerG 685
Db 23173 CAACAAGCGGATACGCGCGCGGAAACGCTT-----CCTCGACGCGCTG 23217
QY 685 lAlaMetLeuGlyLeuLysValGlyGlyMetThrGluSerGlyArgLeu---- 703
Db 23218 GCGACCGCGCGCGAGTCTTGTACAGCGCGACGCTCGCTGTGTGTGTGTGTGTGTGTGT 23277
QY 704 -----CysAlaPheIleThrLysVal 711
Db 23278 GCGCGCGGTGTATATTATTCAGGGGCGGCTGGCGAGCCCGAGCTGGAGCAAGCTGTCTGT 23337

QY 711 yslYsglySerLeuAlaAspThrValGly---HisLeuArgProGly-----AspG 727
Db 23338 CCGATGGCCCTTGGCTGGCGCTGGCGCTGGCGCAAGCCCTGGAGCACGAGAGACC 23397
QY 727 luValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnI 747
Db 23398 ACCGTACCGTCGCGCAGCATGACTGGCGCGCTTTGGCGCTTGCATCAGCGTGGCTCGC 23457
QY 747 le-----IleLeuGluSerLysProGluProGlnVal-----GluLeuValV 761
Db 23458 TCCCGCGCTCTGGCGACTGTCGCGAGCGCGCCCTCGAGACGAGAGCGCGCT 23517
QY 761 alSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGluS 781
Db 23518 CCTCTCCGAGCAGCGCGCGCGCGCGCGCTCTCG-ACAAGCTCCGAGCGCGCTCGGA 23576
QY 781 erSerSerSerPheGluSerGlnLys-----MetAspArgProSerIleSerV 798
Db 23577 GAGCGAGCAGCTCTGCTCTGCGCGCGCTGGTGGCGAGCAGCGCGCTCTGCTCGG 23636
QY 798 alThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnL 818
Db 23637 CCAGAGCGCGCTCCAG-----CTCGACCCCGCAGCAAGCTTCTTCGACCTCGG 23687
QY 818 euSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyA 838
Db 23688 TCTCGAATCGA-----TCATGACCTCGAGCTTCTCGCGG 23723
QY 838 lalysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrP 858
Db 23724 CTTGCAACAGCGCCACCGGCATCAAGCTCCCGGCCACCTCGCCTTCGACCATCTCTCC 23783
QY 858 heLeuProAspArgSerAsp-----864
Db 23784 TCATCGCGCTCGGCTCTTCATGCGCGACTGCTCGCCGACGCCCTCGGACGAGGCTCTC 23843
QY 865 -----LysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLysTrpAsnG 883
Db 23844 CGCGAGCGCGAGCGCGCGGCTCGCGCGCGCTCGAGCGAGCGCCATCGCCATCGT 23903
QY 883 lnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluIle 903
Db 23904 CGGCATGGC--CCTGCGCCTCGCGCGCGCTCGCGATGTCGAGCTCTTTGGAGTTC 23961
QY 903 hrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGluIleLeu- 922
Db 23962 CTCACCAAGGCGCGAGCGGTCGAGCCCATTCACAGAGCGCGCTGGGACGCGGTCG 24021
QY 923 -----IleGluLeuGluThrAlaLeuLeuAsp---931
Db 24022 CTCTAGACCCCGACCCCGAGCGCGAGCCCAAGAGTACTCGCGGATGCGCGATGCTC 24081
QY 932 -----AspGluProHisIleTyrTyrLysLeuGlnThrHisAspValSerLeuProLeuP 950
Db 24082 GACCATGACCTCTTCGACCTCGCTCTTCGCGCATCAGCCCGCGGAGGCCAAACAC 24141
QY 950 roArg---ProSerProTyrLeuProArgArgGlnLeuHisGlyGluSer-----965
Db 24142 CTGACCCCGACCGCGCTGCTCTCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 24201
QY 966 -----ProThrArgArgLeuGlnArg-----SerLysA 975
Db 24202 ATGCTCCCACTCCCTCAAGGACTCTCTCAGCGGCTCTCTGCGCATCTCGCGCGG 24261
QY 975 rgIleSerAspSer-GluValSerAspTyrAspCysGlu-----987
Db 24262 GAATACGCGATCAAGAGCGAGCTCGAAGGTTCCGAGTTTACTTCATCCAGGCAT 24321
QY 988 -----AspGlyValGly-ValValSerAspTyrArgHisAsnGlyAr 1001
Db 24322 TCCGCTCTCTTGGCGCGGCGGCTTGGCTTATACGCTCGGCTCGGCTCGGCTCGG 24375

QY 1001 qAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSe 1021
Db 24376 CGATCTTCGTCGACACCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24435
QY 1021 rProSerGlySerProHisArgValAspValIleGlyArgThr---ArgSerTrpSerPr 1040
Db 24436 GCCCTCCGACAGGCGAGTGCAACTCGCCCTCGCGCGCGGCGTGTCTCATGTCTCTCC 24495
QY 1040 oSerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGl 1060
Db 24496 CCCGAGACCTTCGTCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 24555
QY 1060 yHisTyrAsnThrIle-SerArgMetAspArgHisArgValMetAspAspHisTyrSerS 1080
Db 24556 ACCTTCTCGACACACCGCGAGCTACGAGCGGAGA-----24593
QY 1080 erAspArgAspAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerT 1100
Db 24594 --AGCGCTCTGTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGG 24626
QY 1100 hrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProA 1120
Db 24627 -----CGACGCTCTCGCGCGAGAC-ACCCTCTCTCTCTCTCTCTCTCTCTCTCG 24680
QY 1120 spThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerP 1140
Db 24681 CAACACGAGCGCGCGT-----CGAGCGGTATCACCGCCCC 24716
QY 1140 roAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrP 1160
Db 24717 CAACGCGACCTCCAGCAGAGGTCTCTCGCGCGCGCTCCACGAGCGCGCATCA 24772
QY 1160 roGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProLysGlyThrLeuGluA 1180
Db 24773 -----CCCCCGCGAGCTCGAGCGCT 24794
QY 1180 rgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnV 1200
Db 24795 CGAGTGCATG-----24805
QY 1200 alAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpA 1220
Db 24806 -----GCACCGGCACTCTCTCTGAGACCCCATCGAGGTGCAAG-----CCTGTC 24851
QY 1220 spProHisArg---GlyAlaAspThrValSerThrLysSerSerAspSerAspValSerA 1239
Db 24852 CGCGCTCTACCGCGAGCGAGACCGCTGAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCA 24911
QY 1239 spValSerAlaValSerArgThrSerSerAlaSerArgPhe-----1252
Db 24912 CAACATCGGCGCATCTCGAGCGCGCTCTCGCGCGCTCGCGAGATGTCGCTC 24971
QY 1253 --SerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleS 1272
Db 24972 GCTCGCGCAGCGCT 25016
QY 1272 erValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerLysLysAsnLeuThrL 1292
Db 25017 CAATCCCTCTCATCG-----AGTGGAGGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 25049
QY 1292 ySerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerA 1312
Db 25050 CGACGCTGCTGATACCCGAGGCTT---GGGCC-CCGACGAGATGCGAGTCTCTCTCTCT 25105
QY 1312 spThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgArgSerSerIleGlyAlaL 1332
Db 25106 GCGCGGCT 25165
QY 1332 yeMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeu-----1349
Db 25166 CTCCGCGCGCT 25225
QY 1350 -----SerGlnThrGluGlyGlyLysL 1358

Db 20637 CCTCTCCGCCACGACGACCT-----CGCCCTGCCGATGAGCTA 20678
Qy 41 LeuAlaValMetAspArgGlnLysGluGluGluGluSerValLeuLysile 60
Db 20679 CTGGCT----- 20684
Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 20685 ---CGCCACCAACCCCGGCTA-CCTTCGAGACCGTCCGCTCTCTGGTCCACACCGCG 20740
Qy 81 ValAsn-----AsnValLeuGlnProGlnGlnLysGlnPro----- 92
Db 20741 AAGAGCTCTCTCCCGCTCGATTCTGCTCCGAGGAGGAGCCCGCCCGAGCACCCTCG 20800
Qy 93 ---AsnGluLysGluProGlnThrLysLeuHisGlnGlnPheGluMetTyrLysGluGln 111
Db 20801 TCGAACGAGCGGAGGACACGCGAGGTCGTCTTCTGCTTTTC----- 20842
Qy 112 ValLysLysMetGlyGluGluSerGlnGlnGlnGlnGlnGlnLysGlyAspAlaProThr 131
Db 20843 ---CTGGCC-----AAGCTCCAGTGGAGAGGATGGCCCTCTCCC 20881
Qy 132 CysGlyIleCysHisLysThrLysPheAlaAspGlyCysGlyHisAsnCysSerTyrCys 151
Db 20882 TGCTCGATACCTCCGCGCTCTTC-----GGGCACAGCTCGAAGCGTGG 20926
Qy 152 Gln-----ThrLysPheCysAlaArgCysGlyGly-----Arg 162
Db 20927 ACGCGCCCTCCGCCCCACGTCGACTGCTGCTCGCGTCTCCGCGGAGGAGG 20986
Qy 163 ValSerLeuArgSerAsnLysValMetTyrValCysAsnLeuCysArgLysGlnGlnGlu 182
Db 20987 GCGCGCCCGCTCAGCGGTCGAGTGTGTCCAGCCCGCTGTCTCGATGATGCTCT 21046
Qy 183 IleLeuThrLysSerGlyAla----- 189
Db 21047 CGCTGGCGCCCTGTGGCGCTCCATGGCGTCGAGCCGCGCTGCTCGCCCATAGCC 21106
Qy 190 -----TrpPheTyrAsnSerGlySerAsnThrLeu-GlnG 201
Db 21107 AGGGCGAGATCCCGCGCTGTGTGGCGGCGCGCTGCTCGTCGAGGACGCTCGCAAGC 21166
Qy 201 nProAspGlnLysValPro-----ArgGlyLeuArgAsnGluGluAlaProGlnGlu 219
Db 21167 TGGTGGCGTGGCAGCGCTGCTGTGTGGAGCTCGCGCGCCAGGGGC----- 21215
Qy 219 sLysAlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValPr 239
Db 21216 -----CATGGCGCGGTGGAGCTGCCGAGCGCGAGGTCCGACGCGCGCTCC 21262
Qy 239 oAlaValGluLysGlyArgAlaHisGlyLeuThrArgGln-----AspThrI 255
Db 21263 ACGCTATGGCGATCGGCTCTCCATCGGGGCGATCAACAGCCCTGTTTCAGAGATCT 21322
Qy 255 eLysAsnGly-----SerGlyValLysHisGlnIle-- 265
Db 21323 CGGGCAGAGCCCTCGCGTGGCGCGCTGCTCGCGGATCTGGAGTCCGAGGCGCTTCG 21382
Qy 266 -----AlaSerAspMetProSerAspAr 273
Db 21383 CCTCAAGCTGAGTTACGATTTGCGCTCCATCGGGGCGATCAACAGCCCTGTTTCAGAGATCT 21442
Qy 273 g---LysArgSerProSerValSerArgAsp-----G 283
Db 21443 AGCTCTCGATCTCTCTGCTGGCTCGAGCCGCTCGAGCGGCTCCGCTTCTACTCCA 21502
Qy 283 nAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrValProSerAs 303
Db 21503 CGGTGAGCGCGCGGATCGAGCGGAGCGAGCTCGACGCGCCCTACTGTTACCGGAACC 21562
Qy 303 p----- 303
Db 21563 TCCGGCAGCGGTCGCTTTCGAGACGCTGTGCAAGGCTCTCTTCCGCGAGAACATCGCT 21622

Qy 304 -----GlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArgG 321
Db 21623 TCTTCGTGGAGTGGAGCCCACTCTCTGTCTGACCTTGGCCCTTGCACAGCTCTCTCGAAG 21682
Qy 321 uProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgGlyHis 341
Db 21683 CGTCGAGAGCGCTCGCGCGGCTGTCTCTCTCTG-----GAGCGACGAGGGGATC 21736
Qy 341 s-----ArgHisSerLysGluTyrIleValAs 350
Db 21737 TAGCGCGCT 21796
Qy 350 pAspGluAspVal----- 354
Db 21797 CGACGATCTGTCGCCCCCGGAGAGCGGTCGCTCCACCTACCCCTTCGAGCGCGAGC 21856
Qy 355 -----GluSerAr 357
Db 21857 GCTTCTGGCTCGACGCTCCACGGCACCGCGCGGCGGTCAACCACTTGTCTCGCTCG 21916
Qy 357 gAspGluTyrGluArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspPr 377
Db 21917 AGGGCGGTCTGTGGAGGCGCATCGAGAGCGGAGATATCGACGCGCTCGAGCGCAGCTCC 21976
Qy 377 o-----AsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMe 393
Db 21977 ACGTGGACGCGACGAGCGCGCGCGCTTGC-----CTGTCTCT-----TCCACCCTCG 22030
Qy 393 tArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSerLe 413
Db 22031 CGAGCTTTCGCCACGAGCGCAAGAGAGGCGCACGCTCGACGC-----CT 22075
Qy 413 uAla-----AsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspAr 430
Db 22076 GGGCTTACGCAATCAGTCAAGAGCTCTGACACCGC-----CACACGCGCGCG 22126
Qy 430 gProSerArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTy 450
Db 22127 ACCTGGCGGCGAC-----CTGGCTCTCTGTGTGGCGCGCTTGGAGCAGCGCGC 22180
Qy 450 rSerMetGluArgThrArgGluAlaGln---GlyGlnSerSerTyrProGlnArgThrSe 469
Db 22181 TCCTCTCCCGCTTCCAGAGGCGCTCGCGCGCGCGCGCGCTCTCTCGCGCTCGCGC 22240
Qy 469 rAsnHisSerProProThrPro-----ArgArgSerProIlePro-----Le 483
Db 22241 TGAGCAGCGCCACCTTGGACCGGAGGCTCTCGCGAGCACCTTGGCGAGGCTTGGCGCG 22300
Qy 483 uAspArgProAspMetArgAlaAspSerLeuArg----- 495
Db 22301 AGACCGCGCGCTCGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 22360
Qy 496 -----LysG 497
Db 22361 ACCATCGCGCTGTCCCGCGGACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 22420
Qy 497 nHisHis-----LeuAspProSerSerAlaValArgLysThr-- 509
Db 22421 ACATCGCCCTCGAGCGCGCTTGTGTGCTCTTTCACCGCGCGCGCTCTCTCTCTCTCT 22480
Qy 509 ----- 509
Db 22481 CCGACCCCATCGGCCCATCGAGCGAGGATGATCTGGGCGCTGGCGCGCTGGCTGGCGCG 22540
Qy 510 -----LysArgGluLysMetGluThrMetLeuAr 519
Db 22541 TCGAGCACCCCGAGCGCTGGGAGGCGCTGTGTGAGCTCGCGCGAGCATCGAGCGAGCG 22600
Qy 519 gAsnAspSerLeuSerSerAspGlnSerGluSerValArg-----Pr 533
Db 22601 CGTGGCGCGCTTGTCTCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 22660

Db 24681 CAACACGAGCGCGCT-----CGAGCGGTATCACCGCCCC 24716
Qy 1140 roAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrP 1160
Db 24717 CAACGCGACCTCCAGCAGAGGTCTCCGGCGCGCTCACAGCCCGCATCA----- 24772
Qy 1160 roGlyThrGlyArgGlyArgGlnLeuProGlnLeuProGlnLeuProGlnLeuProGlnLeu 1180
Db 24773 -----CCCCCGCGAGCTCGAGTCTG 24794
Qy 1180 rgSerAlaMetAspIleGluGlnArgGlnMetLysLeuAsnLysTyrLysGlnV 1200
Db 24795 CGAGTGCATG----- 24805
Qy 1200 alAlaGlySerAspProArgLeuGlnAspTyrHisSerLysTyrArgSerGlyTrpA 1220
Db 24806 -----GCACGCGACCTCGCTGGAGACCCATCGAGGTGCAAG-----CCCTGGC 24851
Qy 1220 spProHisArg---GlyAlaAspThrValSerThrLysSerSerAspSerAspValSerA 1239
Db 24852 CGCGTCTACGCGACGCGAGACCGCTGAAAGCCTCTCTCTCGCGCGCTCAAGAC 24911
Qy 1239 spValSerAlaValSerArgThrSerSerAlaSerArgPhe----- 1252
Db 24912 CAACATCGGCATCTCGAGCGCGCTCGCGGCTCGCGGCTGCCAAGATGCTCGCCTC 24971
Qy 1253 --SerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleS 1272
Db 24972 GCTCCGCGACGAGCCCTGCCCCCACCCTCCAGCGAGCCCCAC-----G 25016
Qy 1272 erValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrL 1292
Db 25017 CAATCCCTCATCG-----AGTGGAGGCGCTCGCAT 25049
Qy 1292 ysSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerA 1312
Db 25050 CGACGCTCGTACATCCCGAGCGCTT---GGCGCC-CGCGCAGGAAGTGGCAGTCCCGCC 25105
Qy 1312 spThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgArgSerSerIleGlyAlaL 1332
Db 25106 GCGCGCGCATCTCGCGCTCGATTCCTCGGCGACCAACGCGCCATCATCTCTCGAAGAGG 25165
Qy 1332 ysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeu----- 1349
Db 25166 CTCCGCGCGCTCGCGCGCGAGCCCGACCTCAGACGCGCGCTCGAAGCGCTCCG 25225
Qy 1350 -----SerGlnThrGluGlyGlyLysL 1358
Db 25226 CGCGTGGCGCGTCTCTCGCGCAGGAGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCG 25285
Qy 1358 ysLeuArgSerThrValGlnArgSerThrGlnThrGlyLeuAlaValGluMetArgAsnT 1378
Db 25286 GGCTCCCGACACCTCGTCCGCCACAGCAGCTCACCCCTCGCGGATGTGGCTATTTCG 25345
Qy 1378 rpMetThrArgGlnAla-----SerArgGlnS 1387
Db 25346 TGGCCACACCGCGCGCCCATCTCGAGCAGCGCGCGCTCTGTAGGCCACACCGCGAGC 25405
Qy 1387 erThrAspGlySerMetAsnSerTyrSer----- 1396
Db 25406 AGCTCTCTCGCGCTCGATCGCTCGCCCGCAGGACAGCGCGCGCGCGCGCGCGCGCGCG 25465
Qy 1397 -----SerGluGlyAsnLeu-----IlePheProGlyVal----- 1406
Db 25466 GACGAGCGGAAGCAGCGAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25525
Qy 1406 ----- 1406
Db 25526 AAGGATGGCGCTCTCGCTGCTGACTCTCTGCGCGCTCTCTCGCGCTCTCTCGAGCAT 25585
Qy 1407 -----ArgLeuAlaSerAspSerGlnPheSer----- 1415
Db 25586 GCGAGCGCGCT 25645

Qy 1416 -----AspPheLeuAsp----- 1419
Db 25646 AGGCGCGCGCT 25705
Qy 1420 -----GlyLeuGlyProAlaGlnLeuValGlyArgG 1430
Db 25706 TCTCCCTGGCGCGCT 25764
Qy 1430 lnThrLeuAlaThrProAlaMetGlyAspIleGln-----ValGlyMetMetA 1446
Db 25765 -----AGTCAGGCGAGATCCCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25807
Qy 1446 spLysLysGlyGlnLeuValGluIleLeuArgAlaArgGlyLeuValValLysProG 1466
Db 25808 CCTTCGAGGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25867
Qy 1466 lySerLysThr-----LeuProAlaProTyrValLysValLysValLysValLysLeuA 1481
Db 25868 GCAACGCGCGCATGCGCGCGCTCGAGCTCGCGCGCTCCGAGCTCCAGACTCTCTCTCT 25927
Qy 1481 spAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeu----- 1498
Db 25928 CCTGGGCGGACAGGCTCTCCATCGCGCGCTCAACAGCCCCAGGCGCGCTCTCTCT 25987
Qy 1499 --GluPro-----LeuTyrGlnGlnLeuLeuSerPheGluGluSerProG 1513
Db 25988 GCGAGCGCGCGCGCATGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26037
Qy 1513 lndGlyArgValLeuGlnIleIleValTrpGlyAspTyrGlyArgMetAspHisLysSerP 1533
Db 26038 -----GTCTCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26074
Qy 1533 heMetGlyValAlaGlnIleLeuLeuAspGluLeu-----LeuSerAsnMet-Val 1550
Db 26075 ACTCCGCCAGATGAGCGCGCTCCAGAGAGAGCTCGCGCGCGCTCTAGCCACATCGCT 26134
Qy 1551 lIleGlyTrpPheLysLeuPheProProSerSerLeu---ValAspProThrSerAlaPro 1569
Db 26135 CTGCGA-----CGTGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26179
Qy 1570 LeuThrArgArgAlaSer-----GlnSerSerLeuGluSerSerThrGlyProSer 1586
Db 26180 GGCTCGAGCGCTCGAGCTCGAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26239
Qy 1587 TyrSerArgSer 1590
Db 26240 TGTCTCGAGCG 26251

RESULT 14
US-08-457-335A-6
; Sequence 6, Application US/08457335A
; Patent No. 5723759
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hamner, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532

QY	377	O-----AsnLeuAlaArgTyrProValIysProGlnProTyrGluGluInMe	393
DB	21977	ACGTGGAGCGGCAGCAGCAGCGCCCTTCG---CTGCTCT---TCCCACCCTCG	22030
QY	393	tArgIleHisalaGluValSerArgAlaArgHisGluArgHisSerAspValSerLe	413
DB	22031	CGAGCTTTTCGCCACAGCGCGCAAGAGCAGGCGCATCGTTCGACGC	22075
QY	413	uAla-----AsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspAr	430
DB	22076	GCGCTACCGCATCACGTGGAAAGCTCTGACACCGC-----CACACGCGCGCGC	22126
QY	430	gProSerArgGlnArgSerValSerGluArgAlaAlaMetGluAsnGlnArgSerty	450
DB	22127	ACCTGGCGCGCAC-----CTGGCTCTCGTCTGCTGGCGCGCTCTGGACGACGCGC	22180
QY	450	rSerMetGluArgThrArgGluAlaGln---GlyGlnSerSerTyrProGlnArgThrSe	469
DB	22181	TCCCTTCGCGCTCACCGAGCGCTCGCGCGCGCGCGCGCGCGCTCTCGCGTGCGCC	22240
QY	469	rAsnHisSerProProThrPro-----ArgArgSerProIlePro-----Le	483
DB	22241	TGAGCCAGGCCCACTGACCGCGAGGCTCTGCGCGAGCACCTTCGCGCCAGCTTGC	22300
QY	483	uAspArgProAspMetArgArgAlaAspSerLeuArg-----	495
DB	22301	AGACCGCGCGCTCGCGCGCGCTCTGCTCTGCTCTGCGCCCTCGACGAAGTCCCT	22360
QY	495	-----LysGI	497
DB	22361	ACCATGCGCGTCCCGCGCGACTCGCTTCTGCTCACCTCGTCCAAGCCTCGCGC	22420
QY	497	nHisHis-----LeuAspProSerSerAlaValArgIlysThr--	509
DB	22421	ACATCGCCTCGACGCGCCTTGTGGCTCTTACCCGCGCGCGCTCTCGTTCGACACT	22480
QY	509	-----	509
DB	22481	CCGACCCCATGCCCCATCCAGCGAGCGATGACCTGGGGCGCTGGCGCGCTCGCGC	22540
QY	510	-----LysArgGluLysMetGluThrMetLeuAr	519
DB	22541	TCGAGCACCCGAGCGCTGGAGGGCTCTGACGTCGCGCGAGCATCGACGCGAGCG	22600
QY	519	gAsnAspSerLeuSerSerAspGlnSerValArg-----Pr	533
DB	22601	CGGTGGCGCTTGTCTCGCGTCTCGCTCGCTCGCGCAAGATGAGGACCGCTCTCTC	22660
QY	533	oPro-----ProProArgProHisLysSerLysGlyGlyLysMe	547
DB	22661	GCCCGCGCGGTCTTACGCTCGCGCTCTGCGCGCTCGCGCTCGCG---CGACG	22711
QY	547	tArgGlnValSerLeuSerSerGluGluGluLeuAlaSerThrProGluTyrThr--	566
DB	22712	CGCGCGCGCACGTACTTCAAGCCCGAGGCGACCTCTCTCATCACCGAGGCGCAC	22770
QY	567	-----SerCysAspAspValGluLeuGluSerGluSe	577
DB	22771	GCGCTGGCGCTCACTGCGCGCATGGCTCGCTCGAGAGGCGCGACGACCTCTGCTC	22830
QY	577	rValSerGluLysGlyAspSerGlnLysGlyLysArgLysThr-----	591
DB	22831	ATCAGCGCGAGGCGCGCAGGCGCGCGCTCGAGCTCCACCGCGAGCTCACCGGC	22890
QY	592	-----SerGluGlnGlyValLeuSerAspSe	600
DB	22891	CTGGCGCGCGCGCTCACTTTCGCGCGGTGTGTGTGCGCGACAGGAGCGCTGTG	22950
QY	600	rAnthrArgSerGluArgGlnLysLysArg-----MetTyrTy	613
DB	22951	CTTCTGAGAGCTCGACGCGCGAAGGCTGCGAGGTCGCGCGCGCGCGCGC	23010
QY	613	rgLyGlyHisSerLeuGlu-----AspLeuGluTrpSerGluProGlnIleLys---	630


```
QY 1387 erThrAspGlySerMetAanSerTyrSer----- 1396
Db 25406 AGTCCTCTCTCGCGCTCGACTCGCTCCGAGCAAGCCGCGCCGAGCACCCTCTCG 25465
QY 1397 -----SerGluGlyAanLeu-----IlePheProGlyVal----- 1406
Db 25466 GACGAGCGGAAGCAAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25525
QY 1406 ----- 1406
Db 25526 AAGGATGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25585
QY 1407 -----ArgLeuAlaSerAspSerGlnPheSer----- 1415
Db 25586 GCGAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25645
QY 1416 -----AspPheLeuAsp----- 1419
Db 25646 AGGCGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25705
QY 1420 -----GlyLeuGlyProAlaGlnLeuValGlyArgG 1430
Db 25706 TCTCCCTGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25764
QY 1430 lnThrLeuAlaThrProAlaMetGlyAspIleGln-----ValGlyMetMetA 1446
Db 25765 -----AGTCAGGCGGAGATCGCGCGCCCTCTCTCTCTCTCTCTCTCTCTCT 25807
QY 1446 splysylsGlyGlnLeuGluValGluIleIleArgAlaArgGlyLeuValValLysProG 1466
Db 25808 CCTCGAGGACGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25867
QY 1466 lySerlysthr-----LeuProAlaProTyrValLysValTyrLeuLeuA 1481
Db 25868 GCAACGGGCGCATGGCGCGCTCGAGCTCGCGCGCTCTCTCTCTCTCTCTCTCTCTCT 25927
QY 1481 spAanGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeu----- 1498
Db 25928 CTTGGGGGACAGGCTCTCCATCGCGCGCTCAACAGCCCGGCGCCAGCTCTCTCTCT 25987
QY 1499 --GluPro-----LeuTyrGlnGlnLeuSerPheGluGluSerProG 1513
Db 25988 GCGAGCCCGCGCATCGAGCGGCTGATCGACTCGCTCACCGCGCGCAG----- 26037
QY 1513 lnGlyArgValLeuGlnIleIleValTyrGlyAspTyrGlyArgMetAspHisLysSerP 1533
Db 26038 -----GTCTTCGCGCGAAAGTCCGCGCTCGACTAGCGCTCTCTCTCTCTCT 26074
QY 1533 heMetGlyValAlaGlnIleLeuLeuAspGluLeuGlu-----LeuSerAsnMet-Val 1550
Db 26075 ACTCCGCCGAGATGGAGCGCGTCCAGAGAGCTCGCGCGAGGTCTAGCCACATCTGCTC 26134
QY 1551 IleGlyTyrPheLysLeuPheProProSerSerLeu---ValAspProThrSerAlaPro 1569
Db 26135 CTCGGA-----CGTGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26179
QY 1570 LeuThrArgArgAlaSer-----GlnSerSerLeuGluSerSerThrGlyProSer 1586
Db 26180 GGTCTGACGGCTCCGAGCTCGAGCGCGCTACTGGTATCGAAACCTCTCGGCAACCGCTC 26239
QY 1587 TyrSerArgSer 1590
Db 26240 TGTTCTCGAGCG 26251
```

Search completed: November 16, 2003, 15:18:59
Job time : 992 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 06:43:01 ; Search time 850 Seconds
(without alignments)
5049.537 Million cell updates/sec

Title: US-09-617-099B-1

Perfect score: 8285

Sequence: 1 MSAPLGRGRPAAPTAASQP.....TTRASQSSLESSTGSPSYRS 1590

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USPRO.spool/US09617099/runat_14112003_191545_23673/app_query.fasta_1.1735
-DB=N Geneseq 19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09617099 -CGN 1 1 0 @runat_14112003_191545_23673 -NCPUP=6 -ICPU=3
-NO_MMAP -LARGQUERY -NGS_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8285	100.0	4980	22	Mouse Rim2-encodin
2	2936.5	15.5	6637	22	Human neuroblastom
3	1781.5	21.5	1302	19	Human secreted pro
4	1042	12.6	7325	24	Human DNA sequence
5	1015.5	12.3	2276	21	Human secreted pro
6	992	12.0	7395	23	Drosophila melanog
7	598.5	7.2	526	22	Human nervous syst
8	470.5	5.7	20604	23	Drosophila melanog
9	420	5.1	378	21	Human ORFX ORF489
10	374	4.5	9027	22	Human ORFX ORF489
11	368.5	4.4	24387	22	Human ORFX ORF489
12	368	4.4	355	20	Human ORFX ORF489
13	368	4.4	355	20	Human ORFX ORF489
14	360.5	4.4	9551	20	Human ORFX ORF489
15	311	3.8	6755	19	Human ORFX ORF489
16	298	3.6	16857	23	Human ORFX ORF489
17	298	3.6	18439	23	Human ORFX ORF489
18	298	3.6	40766	23	Human ORFX ORF489
19	298	3.6	42429	23	Human ORFX ORF489
20	296	3.6	9807	23	Human ORFX ORF489
21	295	3.6	7568	23	Human ORFX ORF489
22	295	3.6	9220	23	Human ORFX ORF489
23	295	3.6	9373	23	Human ORFX ORF489
24	295	3.6	9373	23	Human ORFX ORF489
25	295	3.6	9416	24	Human ORFX ORF489
26	295	3.6	12505	23	Human ORFX ORF489
27	294	3.5	5502	23	Human ORFX ORF489
28	293	3.5	4509	23	Human ORFX ORF489
29	293	3.5	4810	23	Human ORFX ORF489
30	292.5	3.5	7796	23	Human ORFX ORF489
31	292.5	3.5	9880	21	Human ORFX ORF489
32	292.5	3.5	23914	23	Human ORFX ORF489
33	292.5	3.5	109519	22	Human ORFX ORF489
34	291	3.5	13095	25	Human ORFX ORF489
35	291	3.5	13152	25	Human ORFX ORF489
36	291	3.5	14328	25	Human ORFX ORF489
37	291	3.5	2162598	25	Human ORFX ORF489
38	290.5	3.5	5796	23	Human ORFX ORF489
39	290	3.5	4483	19	Human ORFX ORF489
40	290	3.5	48300	22	Human ORFX ORF489
41	288.5	3.5	33529	23	Human ORFX ORF489
42	285.5	3.4	6307	23	Human ORFX ORF489
43	284.5	3.4	3819	24	Human ORFX ORF489
44	284.5	3.4	4215	25	Human ORFX ORF489
45	281.5	3.4	18506	23	Human ORFX ORF489

ALIGNMENTS

RESULT 1

- AAF86684
- ID AAF86684 standard; cDNA; 4980 BP.
- XX AAF86684;
- XX AC
- XX AAF86684;
- DT 25-JUL-2001 (first entry)
- XX Mouse Rim2-encoding cDNA.
- XX Mouse; murine; Rim2; Rim isoform; Rab3-interacting molecule;
- XX Rab3-dependent intracellular vesicle fusion regulator;
- XX GTP/GDP exchange factor; CAMP-GEFII; neuron; endocrine tissue;
- XX exocytosis; hormone secretion; neurotransmitter release;
- XX presynaptic membrane; neuropathy; central nervous system disorder;
- XX brain disorder; secretion disorder; pituitary gland; hypothalamus;
- XX pancreatic beta-cell; diabetes; parotid gland; zinc finger; PDZ domain;

KW C2 domain; ss.
XX
OS Mus musculus.

XX Key Location/Qualifiers
XX CDS 32..4804
XX /*tag= a

XX /product= "Mouse Rim 2"
XX /note= "Rim 2 is a novel isoform of Rim"

PN EF1090986-A1.

XX 11-APR-2001.

XX 31-JUL-2000; 2000EP-0116148.

XX 08-OCT-1999; 99JP-0288372.

XX (JCRP-) JCR PHARM CO LTD.

XX (SEIN/) SEINO S.

XX Seino S, Shibasaki T, Ozaki N;

XX WPI; 2001-275168/29.

XX P-PSDB; AAB73488.

XX New Rim 2 proteins and nucleic acids encoding the protein useful in
XX diagnosing endocrine-related diseases or neuropathy and in developing
XX agents for the prevention or treatment of such diseases

XX Claim 4; Page 19-31; 42pp; English.

XX This sequence represents cDNA encoding mouse Rim2. Rim2 is a novel
XX isoform of Rim (Rab3-interacting molecule, also referred to in the
XX specification as Rim1), and both proteins contain a zinc finger,
XX PDZ and two C2 domains. Rim2 is able to specifically interact with
XX the GTP/GDP exchange factor cAMP-GEFII, a cAMP sensor molecule, and
XX is believed to be a regulator of Rab3-dependent intracellular vesicle
XX fusion in neurons and endocrine tissue. It is therefore likely to
XX play a key role in neurotransmitter release at presynaptic membranes,
XX and in hormonal secretion. The Rim2 protein is useful in diagnosing
XX endocrine-related diseases or neuropathy and in developing agents for
XX the prevention or treatment of such diseases. Rim2-specific
XX hybridisation probes or antibodies are useful in the diagnosis of brain
XX or central nervous system disorders, or for diagnosis of secretion
XX disorders of the pituitary gland, hypothalamus, pancreatic beta-cells
XX and parotid gland.

XX SQ Sequence 4980 BP; 1503 A; 1161 C; 1282 G; 1034 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 4980
Score: 8285.00 Matches: 1590
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-617-099B-1 (1-1590) x AAF86684 (1-4980)

QY 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
DB 32 ARGTCGGCTCCGCTCGGGCCCGGGCCCGGGCTCCACCCCGCGCCCTCTCAACCT 91
QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIleIle 40
DB 92 CCTCCGACCCGAGATGCGGACCTCAGCCACCTCAGGAGAGAGAGAGAAATCATC 151
QY 41 LeuAlaValMetAspArgGlnLysLysGluGluGluGlnSerValLeuLysIle 60
DB 152 CTGGCTGTCTATGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGATC 211
QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80

DB 212 AAAGAGAACACAAAGACACACACAGTGGTTCCTTTAGTGGATCACTGACTG 271
QY 81 ValAenAenValLeuGlnProGlnGlnLysGlnProAenGluLysGluProGlnThrLys 100
DB 272 GTAATAACGTTCTGACGCCCCAGCAAAACCAACCAATAGAGAGAGAGAGAGAG 331
QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGlnSerGln 120
DB 332 CTGCACCAACCAATTTGAATGTATAGAGAGAGAGTCAAGAGAGATGGGAGAGATCGCAG 391
QY 121 GlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
DB 392 CAGCAGCAAG 451
QY 141 AlaAspGlyCysGlyHisAenCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
DB 452 GCAGATGATGATCGGCCATAATTTGTTCTATTGCCAACCAAGTCTGTGCTCGATGTGA 511
QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAenLeuCysArgLysGln 180
DB 512 GGTGAGTGTCTTTACGCTCAACCAAGGTATATGTGGGTGTGTAAATTTGTGCCGAAACAA 571
QY 181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAenSerGlySerAenThrLeuGln 200
DB 572 CAAGAAATCTCTCACTAAATCAGAGCATGGTTTTATAATAGTGGGTCTTAACACACTGCAG 631
QY 201 GlnProAspGlnLysValProArgGlyLeuArgAenGluAlaProGlnGlnLysLys 220
DB 632 CAACCTGATCAAAAGGTTCTCTGAGGGCTTCGAATGAGGAGAGAGAGAGAGAG 691
QY 221 AlaLysLeuHisGluGlnProGlnPheGlnIleAlaProGlyAspLeuSerValProAla 240
DB 692 GCAAAACATACACAG 751
QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAenGlySerGly 260
DB 752 GTTGAGAAAGCCGAGCTCTGGGCTCAACAGAGAGATCTATTAAATAGATCAGAG 811
QY 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
DB 812 GTGAAGCACCATGATGTCAGTGCATCTTCAGACAGAAACGAAGTCCATCAGTGC 871
QY 281 ArgAspGlnAenArgArgTyrGluGlnSerGluArgGluAspTyrSerGlnTyrVal 300
DB 872 AGGGATCAAAATCGAAGATACGAGCAAAAGTGAAGAGAGAGAGAGAGAGAG 931
QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320
DB 932 CCTTCAGATGTGTACAAATGCCAGATCTCTTCGGATTATGCTGATAGAGATCTCAGCGT 991
QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAenTyrArgAspSerAsnArgArgGly 340
DB 992 GAGCTCTCAATTTATGAAGAACCTGTGTCTATTAAATTACAGGATTTCTTAACAGAGAGAG 1051
QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
DB 1052 CATAGACATTCCAAAGAGTATATGTTGGATGATGATGAGATGTGAGAGAGAGAGAT 1111
QY 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAenLeuAla 380
DB 1112 GAAGACAAAGAGAGAGAGAGAGAGATACCGAGCACCTACAGAGAGATGATCCAAATCTGCC 1171
QY 381 ArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSer 400
DB 1172 CGGTATCCGTAAGACACACACCTACAGAGACAAATGGCATCCAGCTGAGGTGTCC 1231
QY 401 ArgAlaArgHisGluArgHisSerAspValSerLeuAlaAenAlaGluLeuGluAsp 420
DB 1232 AGGGCAGCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1291
QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440

Db 1292 TCCAGGATTCTCTGCTAAGGATGGATAGACCATCAAGGCAGAAAGATCTGTATCTGAACGT 1351
Qy 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
Db 1352 AGAGCTGCATGGAAGAACCAACCATGCTATTCATGGAAGAACTCGAGAGGCTCAGGGA 1411
Qy 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
Db 1412 CAAAGTTCTTATCCACAAAGGACCTCAATCATAGTCTCTCCACCCCTCGGCGAGCCCT 1471
Qy 481 IleProLeuAspArgProAspMetArgAlaAspSerLeuArgLysGlnHisLeu 500
Db 1472 ATACCCCTTGATAGACAGACATAGCGCGCTGACTCTCTACGGAACAGCACCACTTA 1531
Qy 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
Db 1532 GATCCAGCTCTCTCTGTAGGAAACGAGCGGAGAAAATGGAACCATGTTAAGGAAT 1591
Qy 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
Db 1592 GATCTTTGAGTTTCAGACAGTCCGAGTCAGTGAGCGCCGCCCCACCAAGGCTCATAA 1651
Qy 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAla 560
Db 1652 TCCAGGAAAGGAGGTAAATCGCCAGGTTTCACTGAGCAGCTCGGAGGAGGAGCTGGCA 1711
Qy 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
Db 1712 TCCACACCTGAGTATCAAGCTGTGATGATGTGGAGCTGGAAAGCAGAGGTGTGAGTGAG 1771
Qy 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1772 AAAGGGGACAGTCARAAAGGGAAGAAAGAACTAGTGAGCAGGAGTTTTGTTCGGATTCT 1831
Qy 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1832 AACACCAAGCTCTGAGAGACAAAGAAAGAGATGACTATGTGTGGCCACTCTTTGGAGAG 1891
Qy 621 AspLeuGluTrpSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 540
Db 1892 GATTTGGATGGTCTCAGAGCTCAGATTAAAGGACTCTGGGGTAGATACCTGTAGTAGACA 1951
Qy 641 ThrLeuAspGluGluHisSerHisSerAspLysHisProValTrpGlnProSerLys 660
Db 1952 ACCCTTAAAGGAGAGCATAGCCATAGTATAGCACCCCTGTGACCTGGCAGCCATCCAA 2011
Qy 661 AspGlyAspArgLeuIleGlyArgLysLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 2012 GATGGAGATCGCCTAATTGGTGTGATTTTATTAAATAAGCGTTTAAAGATGGGAGTGA 2071
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 2072 CTTGAGACTCAGGAGCAATGTCTGGGCTTAAAGGTTGTAGGAGAAAGATGACTGAATCA 2131
Qy 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 2132 GTTGCATTTGTGCAATTTATACCAAGTAAAGGAAAGGAGTTTACTGATACCTGAGGA 2191
Qy 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 2192 CATCTTAGACCAAGTGTAGTCTTGGATGGATGGAGGCTATTGACAGGAGCCACA 2251
Qy 741 PheGluGluValTyrAsnIleLeuLeuSerLysProGluProGlnValGluLeuVal 760
Db 2252 TTTGAGGAAGTTTACAACTATTCTAGAAATCCAAACCTGAACCAACAGTTGAGCTTGT 2311
Qy 761 ValSerArgProIleGlyAspIleProArgLysProAspSerThrHisAlaGlnLeuGlu 780
Db 2312 GTTTCAAGGCCAATTGGAGATATTCTAGAAATACCTGATAGCAGCATGCACAACTGGAA 2371
Qy 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2372 TCCAGTTCTAGCTCATTTGAAATCTCAAAAATGGACCGCTCTTCTATATACCTTACTCA 2431

Qy 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2432 CCCATGAGTCTCTGGCATGCTGAGGATGTCGCCGAGTTCTTATCTGGACAGCTTTCATA 2491
Qy 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2492 AAATATGTTTGACAAAGTTGTCACAGTTGATAGTTACAAATTTTGGAGCAAGGAT 2551
Qy 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2552 CTCCTCTCCAGGAAAGATGGGAGGCCAGGAATCTTATGTTAAGATTTACTTCTTCCA 2611
Qy 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2612 GATAGAGTGATAAATAAGAGAGAAACAAACAGTCAGAAACCTTTGGAAACCCAA 2671
Qy 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2672 TGGAACCAAGCTTTCATTTATCTCTGTCCACCGAAGAGAAATTCGTGAAAGATGCTG 2731
Qy 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGlu 920
Db 2732 GAAATTTACCTTTGGGATCAAGCTAGAGTTTCGAGAAGAGAGAGCGAATTCCTAGGAG 2791
Qy 921 IleLeuIleGluLeuGluThrAlaLeuAspAspGluProHisTyrTyrLysLeuGln 940
Db 2792 ATTTAATTTGAATTGGAAACAGCTTTGCTAGATGATGAGCCGACCTGGTATTAAGCTGCAG 2851
Qy 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
Db 2852 ACCCATGATGTCCTCTCATTTGTCATCTCCCTCCCTCTCCCATATCTGCCCGAGGAG 2911
Qy 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgLysSerAspSerGlu 980
Db 2912 CTCCTATGGAGAGAGCCCAACGCGAGGCTGCAAGTCGAGAGAAATTAAGTCACAGTGA 2971
Qy 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisGlnGly 1000
Db 2972 GTGCTCTGACTACGCTGCGAGGATGCGTGGGAGTAGTGTGCAGATTTATGCACAAATGC 3031
Qy 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
Db 3032 CGCGATCTTCAAGGCTCCAGCTGTGCGTCCGAGCAAGATCATGTTCATCAATCATTCG 3091
Qy 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
Db 3092 TCACCATCAGGCTCTCTCTCATCGAGTAGATGTTATAGGAAGCAGAGTTCATGTCGCT 3151
Qy 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 3152 AGTGGCCCTCTCTCTCAAGGAAATGGAAACAGGGGCGACCGAGGAGCACGTCCTACTGC 3211
Qy 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 3212 CATTAACAACAAATTAGCCGAATGATAGACACCGTGTATGATGATGACCATCTACTTCCA 3271
Qy 1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
Db 3272 GATAGAGCAGGATTTGAGACGACAGATAGAGCCCATATCAAGATCCAGATCCAC 3331
Qy 1101 GluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAsp 1120
Db 3332 GAACCAACGGCTCTCTTAGAGCGGACCAACCCGCTCCAGATCTCTGAAAGCTCTCTGAT 3391
Qy 1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
Db 3392 ACAACACCTCATGAGTGTGCTTCTATTAATGACTGGAAGATCTGCCCTCTCTCACCT 3451
Qy 1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
Db 3452 GCCTTATCGAGGTCTCACCCCTGTAACCGGTCTGTCCAGAACGCCCATCAAGTACTCG 3511

QY 1161 GlyThrGlyArgGlyArgGlnLeuProGlnLeuProProlyGlyThrLeuGluArg 1180
DB 3512 GGAACAGGACGAAGGGCGGACAGCTTCCACAGCTTCCACCAAGGGAACATTGGAGAGA 3571
QY 1181 SerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysValGlnVal 1200
DB 3572 AGTGCATATGATATAGAGGAGAGAAATCGCAAAATGAACCTTAAACAAATCAACAGGTA 3631
QY 1201 AlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTyrAsp 1220
DB 3632 GCCGGATCAGACCCAGACTGGAGCAAGATTACCATTCGAAGTATCGCTCAGGATGGAT 3691
QY 1221 ProHisArgGlyAlaAlaThrValSerThrLysSerSerAspSerAspValSerAspVal 1240
DB 3692 CCACATAGAGGGGACAGTACTGTTTCCACTAAATTCCTCGACAGTGTGATGATGATGTA 3751
QY 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerThrSerThrMetSerVal 1260
DB 3752 TCTGCGGTTCACAGGACTAGTAGTCTTCGCTTCAGCAGCACAGCTACATGTCGTC 3811
QY 1261 GlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsn 1280
DB 3812 CATTCAGAGCGCGGAGGAGAAACAGGAAATCAGTGTCTTTACATCCAAATGCAAAAC 3871
QY 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerLysMetGlnAsn 1300
DB 3872 AGACAGATGGCGTGTGGGGAAGAACTTGACCAAAAGCACCATCATCGTGGAGACATG 3931
QY 1301 CysSerLeuGlnLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320
DB 3932 TGCTCATCTGGAGAAAGATGACGGCAGCCAGCTCCGACACTGCGATGGCGGCTGGGTACC 3991
QY 1321 SerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
DB 3992 AGTGGCAAGAGCGCGCATCTAGCATTTGGGGCAAAATGGTAGCTATTGTTGCTCTCA 4051
QY 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArg 1360
DB 4052 CGGAAAGTGCAGTGCCTCTCAACTCAGCCAAACCGAAGGAGGAGGTAAAGACTACGG 4111
QY 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
DB 4112 AGCATCTGTTACAGAGAAGCACGGAGACCGGGTACGAGTGGAGATGAGAACTGGATGACC 4171
QY 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGlyLysAsn 1400
DB 4172 CGCCAGGCGACCGGGATCCACAGATGCGAGCATGACAGCTATAGCTCGGAAGGAAT 4231
QY 1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
DB 4232 CTGATCTTCCCTGGGGTCCGCTCGCCTCTGACAGCCAGTTCAGTGTATTCCTGGATGGC 4291
QY 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
DB 4292 CTGGGCTCTGCTCAGCTAGTGGAGCGCCAGACCCCTGGCTACTCTCTGCATGGGTGACATT 4351
QY 1441 GlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArgGly 1460
DB 4352 CAGGTGGGAGATGATGATATAAAGGAGACAGCTGGAGGTAGAAATCATCGGCGCGCGGC 4411
QY 1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu 1480
DB 4412 CTTGTGGTAAACACAGGTTCCAAAGACACTGCCAGACACCGTATGTCAAGGTGTATCTGTTA 4471
QY 1481 AspAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeuGluPro 1500
DB 4472 GACAAAGGAGTCTGATAGTACCAAAAGAAACCAAGGTGGCGAGAAAGACCTCGAGGCC 4531
QY 1501 LeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleIle 1520
DB 4532 CTGTACACAGAGCTCTTGTCTTCGAGGAGAGGCCCGGAGGGGTGTACAGATCAT 4591
QY 1521 ValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540

DB 4592 GTCTGGGAGATTATGTCGTATGATGATCACAATCCTTTATGGAGTGGCCAGATATC 4651
QY 1541 LeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPheProSer 1560
DB 4652 TTAGATGAATCGAACTATCCAACTATGGTATGGATGGTTCCTCTCTCTCTCTCT 4711
QY 1561 SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGlu 1580
DB 4712 TCCCTAGTAGATCCAACTCGGACCTCTGACAGAGAGCTTCCCATCTCTCTGGA 4771
QY 1581 SerSerThrGlyProSerTyrSerArgSer 1590
DB 4772 AGTCTACCGGACCTCTTACTCTGTTCA 4801
RESULT 2
AAI98097
ID AAI98097 standard; cDNA; 6637 BP.
XX
AC AAI98097;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human neuroblastoma expressed polynucleotide SEQ ID NO 40.
XX
KW Human; neuroblastoma; ss.
XX
OS Homo sapiens.
XX
PN WO200166733-A1.
XX
PD 13-SEP-2001.
XX
PF 02-MAR-2001; 2001WO-JP01631.
XX
PR 07-MAR-2000; 2000JP-0159195.
PR 12-MAY-2000; 2000JP-0140387.
XX
PA (CHIB-) CHIBA PREPECTURE.
PA (HISM) HISAMITSU PHARM CO LTD.
XX
PI Nakagawara A;
XX
DR WPI; 2001-602630/68.
XX
PT Nucleic acids for prognosis of human neuroblastoma comprise nucleic
PT acids expressed by human neuroblastomas -
XX
PS Claim 1; Page 97-101; 159pp; Japanese.
XX
CC The invention relates to nucleic acids (AAI98058-AAI98161) or their
CC homologues expressed by human neuroblastomas useful for detecting genes
CC expressed by neuroblastoma and for analysing their structure and
CC function. The nucleic acids are useful for the diagnosis and prognosis of
CC neuroblastoma.
XX
SQ Sequence 6637 BP; 2020 A; 1423 C; 1511 G; 1683 T; 0 other;
Alignment Scores:
Pred. No.: 1.39e-133 Length: 6637
Score: 2936.50 Matches: 614
Percent Similarity: 66.75% Conservative: 143
Best Local Similarity: 54.14% Mismatches: 232
Query Match: 35.44% Indels: 145
DB: 22 Gaps: 24
US-09-617-099B-1 (1-1590) x AAI98097 (1-6637)
QY 1 MetSerAlaProLeuGlyProArgGlyProAlaProAlaProThrProAlaAlaSerGlnPro 20
DB 321 ATGTCTCGGCGGTGGGCGCGGCTCTCTCGGCGCGCGCGGCTCTCTCTCTCTCTCTCT 368
QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIleIle 40

QY 735 LeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleLeuLeuSerLysProGlu 754
Db 2274 CCCCTGCGGAGCTACAAATGAAGAAGTTTACACATTTATTTAGAAATCAAAATCAGAA 2333
QY 755 ProGlnValGluLeuValSerArgProIleGlyAspIleProAspSer 774
Db 2334 CCTCAAGTTGAATTTATTTGTTCAAGGCTTATGTCATCTCCCGATTCCTGAGAGC 2393
QY 775 ThrHisAlaGlnLeuSerSerSerPheGluSerGlnLysMetAspArgPro 794
Db 2394 TCCCACTCCCTGAGTGGAGTTCAGTTTCAATCTCAGAGATGGAAGGCT 2453
QY 795 SerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGlnPheIleu 814
Db 2454 TCCATTTCTGTATTCTCCACCAAGTCTCGAGCTCTAAAGATGCCCCCAAGTCTTA 2513
QY 815 SerGlyGlnLeuSerIleLeuLeuTrpPheAspLysValGlyHisGlnLeuIleValThr 834
Db 2514 CCAGGGCAACTTTCTGTGAAGTTGTGTATGATTAAGTGGACACACAGCTGATTTGAAT 2573
QY 835 IleLeuGlyAlaLysAspLeuSerArgGluAspGlyArgProArgAsnProTyrVal 854
Db 2574 GTTCTGCAAGCAACAGATCTACCTGCTAGATGAGTGCCTCTCGAAATCCCTATGTA 2633
QY 855 LysIleTyrPheLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLys 874
Db 2634 AAAATGTTATTTCTTCCAGATAGAGTGTATAAAGTAAAGGAGGACCAAAACAGTAAAG 2693
QY 875 LysThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGlu 894
Db 2694 AAATACATAGAACCAAAATGGAATCAACTTTGTCTATTACATGATCATCTGATAGAT 2753
QY 895 PheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluGluGlu 914
Db 2754 TTTAGAGAACCAATGTTAGAAATTAATCTGTGGGACCAACCAAGAGTCAAGAGAGAA 2813
QY 915 SerGluPheLeuGlyGluIleLeuIleGluLeuThrAlaLeuLeuAspAspGluPro 934
Db 2814 AGTGAATTTCTTGGAGAGATCTCATAGAAATGGAGACAGCGCTTTTAGATGATGAACCG 2873
QY 935 HisTrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerPro 954
Db 2874 CATTTGGTATAACTTTCAGACATGATGATCTTCACTACCTCTCCCTCAGCCATCCT 2933
QY 955 TyrLeuProArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLys 974
Db 2934 TTCATGCCCAAGCGACATATTTCTAGAGAAAGCTTAGCAAAAGCTACAAAGATCTCAG 2993
QY 975 ArgIleSerAspSerGluValSerAspTyrAspCysGluAspGlyValGlyValValSer 994
Db 2994 CGAATCAGTATGATGATCTCAGATTTATGAGTTGATGATGATGATGATGATGATGAT 3053
QY 995 -----AspTyrArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSerValProGlu 1012
Db 3054 CAGTAGGCTATAGTCTAGTCTAGAGAAATTAATCTACAACTTAATCTAGTCTAGTCTAG 3113
QY 1013 GlnValMetSerSerAsnHisCysSerProSerGlySerProHisArgValAspValIle 1032
Db 3114 CAGCAAGAACCACTCATCCGCTCAGTCTAGTATCTCTCATCGGCGCAATGATCAG 3173
QY 1033 GlyArgThrArgSerTrpSerProSerAlaProProGlnArgAsnValGluGlnGly 1052
Db 3174 GGAAGCCGCTTACGTTTACCAATGTG---CCATTACAGAGAGTTTATGATGAAT 3230
QY 1053 His-----ArgGlyThrArgAlaThrGlyHisTyrAsnThrIleSerArgMet----- 1068
Db 3231 CATCCAAACAGAGGTCAGCTTCTCCAAACAGACACCATGATGCTCCCGAGTCCAGTT 3290
QY 1069 ---AspArgHisArgValMetAspAspHisTyrSerSerAspArgAspArgCysGlu 1087
Db 3291 GATCATAGAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3350

QY 1088 AlaAlaAspArgGlnProTyrHisArgSerArgSerThrGlu 1101
Db 3351 ATGCTGCCAGACA-----AAACGAGGACGAGTCCAGAA 3386

RESULT 3

AAV40485
XX AAV40485 standard; cDNA; 1302 BP.
XX AAV40485;
XX 09-NOV-1998 (first entry)
XX Human secreted protein CO618_1 cDNA.
XX Secreted protein; CO618_1; human; ds.
XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 123..1247
XX /*tag= a

XX WO9831802-A1.

XX 23-JUL-1998.

XX 21-JAN-1998; 98WO-US01007.

XX 20-JAN-1998; 98US-0010047.

XX 16-JUN-1997; 97US-0072103.

XX 16-JUN-1997; 97US-0877035.

XX (GENY) GENETICS INST INC.

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

XX Racie LA, Spaulding V, Treacy M;

XX WPI; 1998-414101/35.

XX P-PSDB; AAW29640.

XX New isolated polynucleotides and secreted proteins - obtained from
XX human foetal kidney, adult testes, placenta, adult brain and foetal
XX brain cDNA libraries

XX Claim 22; Page 71-72; 104pp; English.

XX This full-length cDNA clone, designated CO618_1 and deposited at
XX ATCC 98296, codes for a novel human secreted protein (see AAW29640).
XX It was isolated from a human adult brain cDNA library using
XX methods which are selective for cDNAs encoding secreted proteins,
XX and was identified as encoding a secreted or transmembrane protein
XX on the basis of computer analysis of the amino acid sequence of the
XX encoded protein. CO618_1 shows some similarity to human male bone
XX marrow myeloblast mRNA for KIAA0237 protein. Novel secreted
XX protein-encoding polynucleotides (see AAV40479-90) have been
XX identified in foetal kidney, adult testis, placenta, adult brain
XX and foetal brain cDNA libraries. These can be used in the
XX production of recombinant polypeptides in host (preferably
XX mammalian) cells. The polypeptides may have e.g. nutritional
XX activity, cytokine and cell proliferation/differentiation activity,
XX immunostimulant or immunosuppressive activity, haematopoiesis
XX regulating activity, tissue growth activity, activin/inhibin
XX activity, chemotactic or chemokinetic activity, haemostatic or
XX thrombolytic activity, receptor/ligand activity, antiinflammatory
XX activity, cadherin or tumour invasion suppressor activity, tumour
XX inhibition activity, or other activities.

XX SQ Sequence 1302 BP; 404 A; 286 C; 322 G; 286 T; 4 other;

Alignment Scores:

Score: 3.5e-78 Length: 1302
1781.50 Matches: 359
Percent Similarity: 88.41% Conservative: 7

Best Local Similarity: 86.71% Mismatches: 7
 Query Match: 21.50% Indels: 41
 DB: 19 Gaps: 1

US-09-617-099B-1 (1-1590) x AAV40485 (1-1302)

1218 GlyTyrAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerVal 1237
 DB 3 GGATGGGATCCTCATAGAGGGGAGATATGTTTCTACTAATCTTCGACAGTGTATGTA 62

1238 SerAspValSerAlaValSerArgThrSerAlaSerArgPheSerThrSerTyr 1257
 DB 63 AGTGATATATCTCGGGTTTCAAGGACTAGTAGTGTCTTCGTTTCAGCAGCACAAGCTAC 122

1258 MetSerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLys 1277
 DB 123 ATGTCTGTCCATCAAGACGCCGAGGAGGAAACAAGAAATCAGTGTCTTTATCATCCAAA 182

1278 MetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSer 1297
 DB 183 ATGCAAGCAGACAAATGGGCATATCAGGAGAGACATGACAAAAGACACGATCATCT 242

1298 GlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAla 1317
 DB 243 CGAGACATGTCTCTCAGAGAGAAATGATGGCAGCCAGTCTGACACTGCAGTGGGCACC 302

1318 LeuGlyThrSerGlyLysArgArgSerSerIleGlyAlaLysMetValAlaIleVal 1337
 DB 303 TTGGGCACCGAGTGGCAAAAGCGCGCTCTAGCCTTGGTGGCCAAAATGTAGTATCGTT 362

1338 GlyLeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLys 1357
 DB 363 GGTCTGTACGGAAGTGCAGTGTCTTCAGTCTAGCCAAACGGAACGAGAGTTAA 422

1358 LysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsn 1377
 DB 423 AAATAGGAGCAGCTCTCCAAAGAGTACAGAAACAGGCGCTGGCCCTGGAAATGAGGAAC 482

1378 TrpMetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerThrSerSer 1397
 DB 483 TGGATGACTCGACAGCAAGCCGAGAGTCTACAGATGGTAGCATGAACAGCTACAGCTCA 542

1398 GluGlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPhe 1417
 DB 543 GAAGGAATCTGATTTCTCGTGTCTGCTTGGCTCTGATAGCCAGTTCAGTGATTTTC 602

1418 LeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMet 1437
 DB 603 CTGGATGGCTTGGCCCTGCTCAGCTAGTGGGAGCCAGACTCTGGCAACACCTGCAATG 662

1438 GlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuValGluIleIleArg 1457
 DB 663 GGTGATTCAGTAGAGATGATGGAACAAAAGGAGCAGCTGGAGGTAGAAATCATCCGG 722

1458 AlaArgGlyLeuValLysProGlySerLysThrLeuProAlaProTyrValLysVal 1477
 DB 723 GCCGTGGCTTGTGTTAAACACAGGTTCCAGACACTGCCAGCACCGTATGTAAAGTG 782

1478 TyrLeuLeuAsnGlyValCys----- 1485
 DB 783 TATCTATTAGATAACGGAGTCTGCTGATGACCAAAAGGAGCAGCTGGAGGTAGAAATCATC 842

1485 ----- 1485
 DB 843 CGGGCCCGTGGCCTTGTGTTAAACACGAGTTCACAGACACTGCCAGCACCGGTATGTAAA 902

1486 ----- 1486
 DB 903 GTGTATCTATTAGATAACGGAGTCTGSAATAGCCAAAGAAACAAAGTGGCAAGAAA 962

1497 ThrLeuGluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgVal 1516
 DB 963 ACCTGGAAACCCCTTTACAGCAGAGTATTATCTCTTTCGAGAGAGTCTCAAGGAAAGTT 1022

QY 1517 LeuGlnIleIleValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyVal 1536
 DB 1023 TTACAGATCATCGTCTGGGAGATTATGCCCATGATCATCAATCTTTATGGAGTG 1082

QY 1537 AlaGlnIleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeu 1556
 DB 1083 GCCCAGATACCTTTAGATGAACCTAGAGCTATCCATATGGTATCGGATGGTTCAACTT 1142

QY 1557 PheProProSerSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGln 1576
 DB 1143 TTCCCACTTCTCTCCTAGTAGATCAACCTTGGCCCTCTGACAAAGAGAGCTTCCCAA 1202

QY 1577 SerSerLeuGluSerSerThrGlyProSerTyrSerArgSer 1590
 DB 1203 TCATCTCGAAAGTTCACTGGACCTTCTTACTTTCGTTC 1244

RESULT 4

AAS94817
 ID AAS94817 standard; DNA; 7325 BP.

XX AAS94817;

DT 14-FEB-2002 (first entry)

XX Human DNA sequence #72 expressed during foam cell differentiation.

XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;
 KW cardiovascular disorder; coronary artery disease; gene therapy; ds.

XX Homo sapiens.

XX WO200177389-A2.

XX 18-OCT-2001.

XX 04-APR-2001; 2001WO-US11128.

XX 05-APR-2000; 2000US-195106P.

XX (INCY-) INCYTE GENOMICS INC.

XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
 PI Tai J;

XX WPI; 2002-010925/01.

XX Composition useful for diagnosis of conditions, disorders or diseases
 PT associated with atherosclerosis, comprises several polynucleotides that
 PT are differentially expressed in foam cell development -

XX Claim 1; Page 123-125; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide
 CC sequences that are differentially expressed during foam cell
 CC differentiation. The polynucleotide sequences of the invention or a
 CC composition comprising these polynucleotides are useful as a high
 CC throughput method for detecting altered expression of one or more
 CC polynucleotides in a sample. The polynucleotides can be used in the
 CC diagnosis of disorders associated with foam cell development such as
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
 CC coronary artery disease. The polynucleotide sequences can also be used
 CC as PCR primers and probes. The polynucleotides of the invention are also
 CC useful in gene therapy. AAS94746-AAS95021 represent the human
 CC polynucleotide sequences of the invention which are differentially
 CC expressed during foam cell differentiation.

XX Sequence 7325 BP; 1526 A; 1985 C; 2033 G; 1778 T; 3 other;

Alignment Scores:

Pred. No.: 1,95e-41 Length: 7325
 Score: 1042.00 Matches: 245
 Percent Similarity: 62.65% Conservative: 62

Best Local Similarity: 50.00% Mismatches: 113
Query Match: 12.58% Indels: 70
DB: 24 Gaps: 13

US-09-617-099b-1 (1-1590) x AAS94817 (1-7325)

```
Qy 1125 ArgSerMetPro-----SerLeuMetThrGlyArgSerAlaProProSerProAla 1141
Db 66 CGGCGCATGCCCGCGTGGCTCCCTC-----GCCGCCCTCCCGCCCGCG 113
Qy 1142 --LeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerThrPro 1160
Db 114 TGCCTATCCACTCGAGTCCGCCGACCTGG-----GCCGGCGCGCTACTGCG 167
Qy 1161 GlyThr-----GlyArgArgGlyArgGlnLeuProGlnLeu-----ProProLysGly 1176
Db 168 GGTTCGGCGGGCGGGTCCCGGGGACGACCTGCGCCGCTTGGAGCGCGCTCG--- 224
Qy 1177 ThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLys 1196
Db 225 CCTGTGGAGGCCCTCCCTG----- 245
Qy 1197 TyrLysGlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArg 1216
Db 246 -----TCTGGACCCCGCGCCACCTCCGACCCCTTTATCATCATGCG 287
Qy 1217 SerGlyTyrAspProHisArgGlyAlaAspThrValSerThrLysSer-----Ser 1233
Db 288 CTCCTCTGGAGCTGCGCTCATTCCTTCACCTCATTTCTTGAATGGTGTGTTTGGC 347
Qy 1234 AspSerAspVal-SerAspValSer-----AlaValSerAspThrSerSe 1248
Db 348 AGAAGTCAATTGAAGCCCTTGCAAAATCCCTAGGGGTGTCTGTGGAGGCGACGCCCC 407
Qy 1248 xAlaSerArgPheSerThrSerTyrMetSerValGlnSerGluArgPro----- 1265
Db 408 TGTGATGGGAACACCGAGCTCAGATTTCATGATCGAGCTGCTGAGGCCCTGCCACCC 467
Qy 1266 -----ArgGlyAsnArgLysIleSerValPhe-ThrSerLysMetGlnA 1280
Db 468 CGGGGACCATGTTTAAACGGGAGCCAGGTCTCTGCTCCTCCTGCGGCTCC----- 517
Qy 1280 snArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspM 1300
Db 518 -----AGGAATGTGTGCGGAGCTCCAGCATTCGCGTGAA 554
Qy 1300 etCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGluAlaLeuGlyT 1320
Db 555 TCTGCGGATCCAGACCGCGGGG-----GGGGCTGGGACA 593
Qy 1320 hrSerGlyLysArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuS 1340
Db 594 CCACCGCAAGAAGCGCGGAGCAGCTGGGTGCCAAGATGGTGGCCATCGTGGGCTGA 653
Qy 1340 erArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyLysLysLeuA 1360
Db 654 CTCAGTGGAGCAAGAGCACATCCAGCTTCGACGCTGAAGGGGCCCAAGAGAGCTGC 713
Qy 1360 rgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetT 1380
Db 714 GCAGCAACATCCCGCGGAGCACGAGACAGGATCGCGGTGGAGATGCGGACCGGGTCA 773
Qy 1380 hrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyA 1400
Db 774 CACGCCAGGCGAGCGGGAGTCCACCGATGGGAGCACCAACAGACGCTCCGACGCA 833
Qy 1400 snLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspG 1420
Db 834 CGTTTCATCTTCCCACTACCCGCTAGGGCTGAAAGCCAGTTCAGCGATTTCCTGGATG 893
Qy 1420 lyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspI 1440
Db 894 GGCTGGGACCACTCAGATTGTGGGCGGACACACTGGCAACACCCATGGGAGATG 953
```

```
Qy 1440 leGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleLeuArgAlaArg 1460
Db 954 TGCACATTGTCATCGACCGGAGTGGCGAGCTGGAGGTGGAAGTGAAGAGCTCGGG 1013
Qy 1460 lyLeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuL 1480
Db 1014 GCTTGACCCCAACAGGCTCCAAATCCCTCCAGCCACCTATATCAAGTTTACTGC 1073
Qy 1480 euAspAsnGlyValCysIleAlaLysLysThrLysValAlaLysLysThrLeuGluP 1500
Db 1074 TGGAGATGGGGCTGCTTGGCCACAGAGAGACAAAGATGACCAAGAGACCTGTGATC 1133
Qy 1500 roLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGluIleI 1520
Db 1134 CCTGTACACAGCAGGCTGTCTCTTTACAGAGGAGCCCGAGGCAATGCTCGAGGTGA 1193
Qy 1520 leValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleL 1540
Db 1194 TCTCTGGGAGACTATGGCGCATGGACACACAGTCTTCATGGGCATGGCCAGATCA 1253
Qy 1540 euLeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPheProProS 1560
Db 1254 TGCTGGAGGAGCTGGACCTCAGCGCGCGGTCCACCGGCTGGTACAAACTCTTCCACCT 1313
Qy 1560 erSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuG 1580
Db 1314 CCTCAGTGGCAGACTCCACACTCGGATCCCTCAGGCGCGCTGTCCCGTCTTCCCTGG 1373
Qy 1580 luSerSerThrGlyProSerTyr-Ser 1588
Db 1374 AGAGTGCCACCGCCCTCATGTCTCT 1399

RESULT 5
AAC60011
ID AAC60011 standard; cDNA; 2276 BP.
XX AAC60011;
XX AAC60011;
XX 26-JAN-2001 (first entry)
XX Human secreted protein gene 46 SEQ ID NO:56.
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antiinflammatory; antitumor; anticonvulsant;
XX antibacterial; antifungal; antiparasitic; cardiact; gene therapy;
XX cancer; immune disorder; cardiovascular disorder; wound healing;
XX neurological disease; infectious disease; chromosome identification;
XX Chromosome 6; ss.
XX Homo sapiens.
XX WO200058356-A1.
XX 05-OCT-2000.
XX 22-MAR-2000; 2000WO-US07535.
XX 26-MAR-1999; 99US-0126511.
XX 17-DEC-1999; 99US-0172413.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-594639/56.
XX P-PSDB; AAB34818.
XX Fifty nucleic acid molecules encoding human secreted proteins, useful
XX in the prevention, treatment and diagnosis of cancer, immune disorders,
XX cardiovascular disorders and neurological diseases -
XX
```

Claim 1; Page 372-373; 425pp; English.

XX The polynucleotide sequences given in AAC59966 to AAC60015 encode the
CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
CC AAB34852 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
CC anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
CC vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for
CC preventing, treating or ameliorating a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The polynucleotides are
CC useful for chromosome identification. They are also useful as probes for
CC diagnosing a disorder related to the female reproductive system,
CC particularly breast and/or ovary cancer. They are also useful in the gene
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC agonists and antagonists from the present invention are useful in the
CC diagnosis, treatment and prevention of cancer, immune disorders,
CC cardiovascular disorders, wound healing, neurological diseases and
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
CC used in the exemplification of the present invention.

XX Sequence 2276 BP; 699 A; 451 C; 436 G; 689 T; 1 other;

Alignment Scores:

Pred. No.: 1,1e-40 Length: 2276
Score: 1015.50 Matches: 197
Percent Similarity: 90.76% Conservative: 19
Best Local Similarity: 82.77% Mismatches: 21
Query Match: 12.26% Indels: 1
DB: 21 Gaps: 1

US-09-617-099B-1 (1-1590) x AAC60011 (1-2276)

QY 1353 GluGlyGlyCysLysLysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAla 1372
DB 84 GAGTCGGCGCCACAAAAGTTAAAGTATCCATCCAGACAGACAGACAGGATGGCA 143
QY 1373 ValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThrAspGlySerMet 1392
DB 144 GCTGAATGAGAAG---ATGGTAAGGACCCGAGCGGAGGCTACTGATGACAGCATC 200
QY 1393 AsnSerTyrSerSerGluGlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSer 1412
DB 201 AACAGTTACAGCTCTGAGGGCAATTTAATATTTCTGGAGTGGGAGCTGACAGT 260
QY 1413 GlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeu 1432
DB 261 CAATTCAGTGATTTCTTGATGATTTGGGACAGCCAGCCAGCTTGTGGCCCAACCCCTT 320
QY 1433 AlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGlu 1452
DB 321 GCCACCCCTCCATGGGTGATATACAAATAGGAATGAGGACAAAGAGGCCAATAGAA 380
QY 1453 ValGluIleLeuArgAlaArgGlyLeuValValLysProGlySerLysThrLeuProAla 1472
DB 381 GTGGAGTCATTAGACGACGAGCCCTCACACAAAGCCTGTTCCAAATCTACACCTGCT 440
QY 1473 ProTyrValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLysLysThrLys 1492
DB 441 CCATATGTCAAAGTATATCTTTTGGAAATGGGCGCTGTATAGCAAGAGAGACAAAGA 500
QY 1493 ValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerPro 1512
DB 501 ATTGCACGAAAAACCCCTTGATCTCTTGTATCAGCAGCTCTGTTTGTATGAAAGTCCA 560
QY 1513 GlnGlyArgValLeuGlnIleLeuValTyrGlyAspTyrGlyArgMetAspHisLysSer 1532
DB 561 CAGGTAAAGTTCTTCAGGTGATTTGCTGGGAGACTATGGCAGATGGACCAATATGC 620

QY 1533 PheMetGlyValAlaGlnIleLeuLeuAspGluLeuLeuSerAsnMetValIleGly 1552
DB 621 TTATGGGGTGGCTGAGATCTTGTGGAGAACTGACCTGTCCAGCATGGTATCGGA 680
QY 1553 TrpPheLysLeuPheProSerSerLeuValAspProThrSerAlaProLeuThrArg 1572
DB 681 TGGTACAAATGTGTCCACCGCTCTCACTGTGGATCCACACTCACTCCCTCACCCG 740
QY 1573 ArgAlaSerGlnSerSerLeuLeuSerThrGlyProSerTyrSerArgSer 1590
DB 741 CGGGCTTCCAGTCATCTCTGAAAAGTTCACTGGGCGCTCCCTGTATTCGATCA 794
RESULT 6
ABL12715
ID ABL12715 standard; cDNA; 7395 BP.
XX ABL12715;
AC ABL12715;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 32627.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI P-PSDB; ABB68612.
DR WPI; 2001-656860/75.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Claim 1; SEQ ID NO 32627; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 7395 BP; 1694 A; 2319 C; 2135 G; 1247 T; 0 other;
SQ Alignment Scores:
Pred. No.: 5.32e-39 Length: 7395
Score: 992.00 Matches: 514
Percent Similarity: 28.11% Conservative: 201
Best Local Similarity: 20.20% Mismatches: 495
Query Match: 11.97% Indels: 1336
DB: 23 Gaps: 72
US-09-617-099B-1 (1-1590) x ABL12715 (1-7395)

1959 CGTCCACGCCCGCCCTCCGCCACCGCCCGTCCAGCGTCCGGCTCCGGTTC 2018
QY 810 ----- 810
2019 CGCCCTGGCGCTGCAAGGCAACAGCGGGGCCACCATCACACCGCCGCCA 2078
Db 810 ----- 810
2079 CCACTACACCCGCCACGAGTCCGCTTCGGATCCGCCACGAGCATCAGGGTCCGCCGC 2138
Db 810 ----- 810
2139 CGCCACCTGCAGGCCACCGGGTGGCGGTGGCATCGGCATCGGCAGCGCTCCGGCAC 2198
QY 811 ----- 811
2199 CACCAACAGCCCATCCCATCGAGGACGGCTGCAGCTGAAGCTTGGCTACACAGAA 2258
Db 811 ----- 811
827 LGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArgGluAspG1 847
2259 CACCCTGCAGCTGATGTAACCTGTGTCTCGGCACCGGACTCTCCCTGCGCCAGAGCG 2318
Db 827 ----- 827
847 YArgProArgAsnProTyrValIysIleTyrPheLeuProAspArgSerAspIysAsnLy 867
2319 AGCGGGCGCAATCCATATGCAAAAGTGTCTCTGCGGCACCGGACTCTCCCTGCGCCAGAGCG 2378
QY 867 ----- 867
867 sArgArgThrLysThrValIysIleTyrPheLeuProLysTyrPheLeuThrPheIleTyr 887
2379 GCGGCAACGAACAGCGTGGGCAACCTCGCAACCCCGGATGGGGCGAGACCTTTGTCTA 2438
Db 887 ----- 887
887 rSerProValHisArgGluPheArgGluMetLeuGluIleThrLeuTrpAspG1 907
2439 CTGGGTCTGAGCGCTGCATCTCAACGCCGACTGTCTCAGGTGACGCTGTGGACTA 2498
QY 907 ----- 907
907 nAlaArgValArgGluGluSerGluPheLeuGlyGluIleLeuIleGluLeuLuth 927
2499 T-----GTGCGGTACGGGGCAACGACTTCATCGCGAGGTGTATCATGATCTGGCGCA 2552
Db 927 ----- 927
927 rAlaLeuLeuAspAspLupProHisTyrTyrLysLeuGlnThrHis 942
2553 CCATCTCTGGACGAGCGCGAGTGTATCAGCTGCAGCTGCAGCCCAACGAGCACCTCTCTA 2612
QY 943 ----- 943
2613 TCTCTTACGTGACGAGGCGAGCGATGTGACGCGCTGATGACTGACCGGACAGATCATTT 2672
Db 949 ----- 949
2673 ATCACCGCGAGCACCATGTGCGCCTTAGTGATTCGGACACCAACGTCGAGCTGGACAT 2732
QY 950 ----- 950
2733 CGATGAATGACTCCCGGGGCCAGCATCTCGTCCATGGGCGAGCTCAGCGAGTCCGCGCC 2792
QY 954 ----- 954
2793 CCGTCTAGAGCTCGATCAACAGAGGTCTGGTCCGACGTGACATGTCGCCCC----- 2844
QY 969 ----- 969
969 gLeuGlnArgSerLysArgIleSerAspSerGluValSerAspTyrAspCysGluAspG1 989
2845 ----CAGGGCGCAACAGGTGGCGGATGGTGGCGCGGACTACCGCACTGTATCTGG 2900
QY 989 ----- 989
989 yValGly-----ValValSerAspTyrArgHisGlnGlyArg 1001
2901 CATTGGACAAAGTTACCAATCAGGCCTCTGCCACCGGCTACTATCTGTCGCGAGGTGG 2960
Db 1001 ----- 1001
1001 gAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSer-----SerAsnHi 1019
2961 TAATGGTGGCTCCGCCATCGGTCCCGGTGGCATGAGCTTCAGCGAGCGGAGTCACTC 3020
QY 1019 ----- 1019
1019 sCysSerProSerGlySerProHisArgValAspValIleGlyArgThr----- 1035

3021 GGCGGCACCCAGCGAGCGTTTACCACAGCAGCGGAGTAGGAGGAGCGGCTCCGGCGGACC 3080
QY 1036 ----- 1036
3081 CGCTAAGGCTATCGGAGCACCGAGTCCCTTGTGCGCCACCCGACGACCG 3140
Db 1047 ----- 1047
1047 g-----AsnValG1 1050
3141 CTACATAGACTATCCAGTCTTCCAGTACACGGCTCTCTCGCTTACGCCCCATCGGTCTA 3200
Db 1050 ----- 1050
1050 uGlnGly----- 1052
3201 CCAGGACCGGGTGTGTGTCTCTCAGCGCGCGCATCTGCTCGCAGACAGCGATTCCA 3260
QY 1052 ----- 1052
3261 GTCCGATCGGCCACAGCCACCGCCACGGGATCTCCAAAGAGAGCGAGCTACACAGGT 3320
Db 1052 ----- 1052
3321 GCGCAGACCTCAGCGAGTGGATGCTGCGAGATCGACTGGGTTCAGGACTTCGACGAGCG 3380
QY 1053 ----- 1053
3381 GCTGGCTCGGGTGTGTCTCGGGCGGCACCGCACACGGCAGCGCCACACCGAGCCAC 3440
QY 1056 ----- 1056
1056 r-----ArgAlaThrGly-----HisArgGlyTh 1069
3441 CTACCGCAGCACCGGAATGGCGGTGGGAGCGCCACTACCGGCTGTCTCGGACAGCGA 3500
Db 1069 ----- 1069
1069 p-----ArgHisArgValMetAspAsp----- 1076
3501 CCTGCACTCTATGGACGCGAGTGCAGCACAAGGCACTCGTTATCCCGCGACAGGACTT 3560
QY 1076 ----- 1076
3561 TATGGCGAGTTCGGCGACTCGGACATGGATGGTGGTGCAGCGTGCAGCTCCAGCGCCTT 3620
Db 1076 ----- 1076
3621 CTCCAGCAGTCCGAGCGACCGCGACGTTCGCGCGGACTCAGCTTCCCGCGCACTGGCG 3680
QY 1076 ----- 1076
3681 CAATCTCTTTGGGTCCGCAACAGCTTCTCCGAGGAGCAGGAGCGCAACCCATTACCGG 3740
QY 1077 ----- 1077
3741 ACTGCGGTGCACCACTCGGAGCC-CGGAAGGCGCAATACCATCGAGGTGGATGACTGCG 3799
Db 1082 ----- 1082
3800 ACTATCTCGCGGTTCGGTGGCGGAGTGGTGTCTGCTGGAGCAGCTGGAGCACTGC 3859
QY 1083 ----- 1083
3860 AGCAGTCCCGCTCTTGGCGAGCGAGTCCCAACCCATTTCCGCGCGGAAATTGGATAC 3919
Db 1086 ----- 1086
3920 AGCCACCTCACCAAGTTTGTGTTACCGATGCCAATAGCGGCCAATTTGGTGGACAACTCT 3979
QY 1087 ----- 1087
3980 AGCTTGAGCCCGGCAATGGCGGAGGAGATGATGGAGCCACTCGATCCCATTTACATC 4039
Db 1093 ----- 1093
4040 CACATTCCCTCTTACTCGCATCGCCCATTTGTCACACAGCAATTCGACATCGATA 4099
QY 1094 ----- 1094
4100 TCGATACCATATATCCACCACCGCGGTCCAAAGGACTGCAACCGAACGTAACGACCGCG 4159

```
QY 1110 ----- 1110
Db 4160 CCGCTGCTATCCGTCGCTGCTGCGAGCTTTGAGCAATTCAGGCGCATTCACGCG 4219
QY 1111 -----
Db 4220 CCATCAGCAATTTCTTCGCGAGCTCCTGCGCGCGGCTCATCCACACGCCACGACC 4279
QY 1121 rAsnLeu-----Me 1124
   :|:|:|
Db 4280 ACACCTCCACAGCGCCAGTTCGCTGGTGGGCTATGTACGTGACCACCGACGACAGGCT 4339
   :|:|:|
QY 1124 tArgSerMetProSerLeuMetThrGlyArgSerAla-----ProProSerPr 1140
   :|:|:| :|:|:|
Db 4340 GCAGCCACTGCAGAAATGGAAGTCAGCGGCTGCTATCGACGCGCACCCACCCCTTCACC 4399
   :|:|:| :|:|:|
QY 1140 oAlaLeuSerArgSerHisProArg-----ThrGlySerValGlnThr 1154
   :|:|:| :|:|:|
Db 4400 TGACCGCGCGGCTCTGCAATCGCGCTTGCAGGATCCCGTGTCTGCGCGGATGCC 4459
   :|:|:| :|:|:|
QY 1155 -----Ser-ProSerSer----- 1158
   :|:|:| :|:|:|
Db 4460 ATCCGACCCACAGCATCAGCGGATGGGCTATGGGAGCCACCTCTTCGGATCCTGGA 4519
   :|:|:| :|:|:|
QY 1158 ----- 1158
Db 4520 TGTGCGCGAGTGTGTGATCAGCACTTCCTGGGTGATTAACGGGACCGCAACTGAATC 4579
   :|:|:| :|:|:|
QY 1159 -----ThrProGlyThrGlyArgGlyA 1167
   :|:|:| :|:|:|
Db 4580 TTGGTGTGTTCCCACTTACCATTGTTCACAGCCCACTCCAAATGCCATCGCGGTGCCA 4639
   :|:|:| :|:|:|
QY 1167 rg-----GlnLeuProGlnLeuProPro----- 1174
   :|:|:| :|:|:|
Db 4640 GGGGCAAAATGGCAACTCTGCACCTTCGCCACCAATCAGTGTACGCTTCGCGCC 4699
   :|:|:| :|:|:|
QY 1174 ----- 1174
Db 4700 AATTGAGCGAGATGCCCTCGTGGCAGCGTTCCTCCATTTCCGAGGCCAAGCGCAACCA 4759
   :|:|:| :|:|:|
QY 1175 -----LysGlyThrLeuGluArgSerAlaMetAspIleGluArgAsnA 1190
   :|:|:| :|:|:|
Db 4760 CTTGATCTCTGAAGAAACCGAGCTGGAGCGAAGCGACTCTTTCACGAGGACAAATCGC 4819
   :|:|:| :|:|:|
QY 1190 rgGlnMetLysLeu----- 1194
   :|:|:| :|:|:|
Db 4820 GGTGCTAGACTACGATGATCTGCACAGAGAGCTGGGTGCGAGCTCGAATTCGTAGC 4879
   :|:|:| :|:|:|
QY 1195 -----AsnLysTyrLysG 1199
   :|:|:| :|:|:|
Db 4880 AGGACGAGGTGATGCTTCGCTGGAGACTATCCGTATCCCATTCGAATCCGTATCCA 4939
   :|:|:| :|:|:|
QY 1199 lnValala-----GlySerAspProArgLeuGlnA 1210
   :|:|:| :|:|:|
Db 4940 GCTACTCACCCACCGCGGCAATGTGCGAGCGTACGGATCGGAGCAATATTCACAAA 4999
   :|:|:| :|:|:|
QY 1210 spTyr-----HisSerLysTyr----- 1215
   :|:|:| :|:|:|
Db 5000 GTTATATGGTGCAAAATGTGGATGCCAATCATCCGCTGAGCCATTCGCAATTTCTGGTAA 5059
   :|:|:| :|:|:|
QY 1216 -----ArgSerGlyTrp----- 1219
   :|:|:| :|:|:|
Db 5060 CGGATGACAAGATAGTGACCAATTACTACGACTCGGATGTGGCTGGAGCGACGCGGGG 5119
   :|:|:| :|:|:|
QY 1219 ----- 1219
Db 5120 TGGCTCCGTCCTCTTTTCGGGAGCGCACCGCAACGCTGGACTGGCGGATCCACGAAATC 5179
   :|:|:| :|:|:|
QY 1219 ----- 1219
Db 5180 ACATGCTGTGATTTGCAGCGAACCCAGCCAAACAAAATTTGTACGACCGCGGCGACCGT 5239
   :|:|:| :|:|:|

QY 1219 ----- 1219
Db 5240 ATTTGAGAGCCCGAAGAAATCTCCGATCCGCCAGTCCGCGAGATGCCAATGCCAATCTTT 5299
QY 1220 -----AspProHisArgGlyAlaAsp----- 1226
   :|:|:| :|:|:|
Db 5300 CACCAATGGAGCCGGATCGGATCGGAATGGATGGCTTAATGCTGCGGACGGGAATGCTGG 5359
   :|:|:| :|:|:|
QY 1227 -----ThrValSerThrLysSerSerA 1234
   :|:|:| :|:|:|
Db 5360 AACAGGGCAACGGAATGGGTAGCGGTGTGCACACACACTACCCAGTACACACAATAGCA 5419
   :|:|:| :|:|:|
QY 1234 spSerAspValSerAspValSerAlaValSerArgThrSerSer----- 1248
   :|:|:| :|:|:|
Db 5420 ACCACCAATATAGCCAAAGCAGTAGTGTAGTGTAGTCCAAAGTCAACAATCAACAATTC 5479
   :|:|:| :|:|:|
QY 1248 ----- 1248
Db 5480 AACAAACAACAATTAACAACAACATCAGTTAGGCAAAAGGAAAGAGGACCAACCG 5539
   :|:|:| :|:|:|
QY 1248 ----- 1248
Db 5540 CCATTAGCGGATCGATGATGGCATCGGAGCAGCTGACGAAATCTAGTAGTGGAGCGGTG 5599
   :|:|:| :|:|:|
QY 1249 -----AlaSerArgp 1252
   :|:|:| :|:|:|
Db 5600 GTGCAACATCTGTCGCGCTGCTGGGCATCTTGTCTGCTGGTGTACTAATGCTAATAATG 5659
   :|:|:| :|:|:|
QY 1252 heSerSerThrSerTyrMetSerValGlnSerGluArgProArg----- 1266
   :|:|:| :|:|:|
Db 5660 CTCTCTCTCTCTTTTGCAAATCCCAACCAACCAACCAACCAACAACAACAACAATCTC 5719
   :|:|:| :|:|:|
QY 1267 -----GlyAsnArgLysIleSerValPheThrSerLysMetG 1279
   :|:|:| :|:|:|
Db 5720 ATGATCACCCCAACAATCTCAACAACAATAACAACGTCGCCCATCATCAACACCAACA 5779
   :|:|:| :|:|:|
QY 1279 lnAsnArgGlnMetGlyValSerGlyLysAsn----- 1289
   :|:|:| :|:|:|
Db 5780 AAAACACGAACACCAATAAATGACGCCAACCAACCTGACGCTGCGACGCCACACAACA 5839
   :|:|:| :|:|:|
QY 1289 ----- 1289
Db 5840 CACAACAACAACAACACTACACAACAACCAACCAATATCATGCANAATCAAAACACGATCCA 5899
   :|:|:| :|:|:|
QY 1290 -----LeuThrLysSerThrSerIleSerGly----- 1298
   :|:|:| :|:|:|
Db 5900 CAACCAATTAACAACAACAACAACAACATCGTGAATGCAACGATGCTGCGACACGA 5959
   :|:|:| :|:|:|
QY 1298 ----- 1298
Db 5960 CAACCGCAATGCAACAATAATGCTACAACAACACTGCCCAAAATCCAAACAACAACAACA 6019
   :|:|:| :|:|:|
QY 1298 ----- 1298
Db 6020 ATAACAAGAACGAACCAACCAACAACAACAACAACAACAACAACAACAACAACAACA 6079
   :|:|:| :|:|:|
QY 1298 ----- 1298
Db 6080 CGGATGCTGACGCGGACGCTGATGCTCGCTGGAGCCCATCGACTGGGATGCCATCGATG 6139
   :|:|:| :|:|:|
QY 1299 -----AspMetCysSerLeuGluLysAsnA 1307
   :|:|:| :|:|:|
Db 6140 CTATGCTGGACGACGACTTTTAGCGAGTACGAACAAGGACAAGACCGCGCCGAAGAGGTG 6199
   :|:|:| :|:|:|
QY 1307 spGlySerGlnSer-----AspThrAlaValGlyAlaLeuGlyT 1320
   :|:|:| :|:|:|
Db 6200 GTGGCTCCAGTACAGCGAGGCTGCGAGTGCAGAGTAAGTGAAGTGCAGGCTCTCTCGG 6259
   :|:|:| :|:|:|
QY 1320 hrSerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeu 1340
   :|:|:| :|:|:|
Db 6260 ACACCGCAACGATCGGAAGAGGCGCGCGC-----GTGACACGAG 6301
   :|:|:| :|:|:|
QY 1340 erArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyLysLysLeuA 1360
   :|:|:| :|:|:|
```



```
Db 123 CGGACCACCCCGCTCCAGATCCACTGACGCTCTGATACAAACCTCATGAGTGGATG 182
QY 1128 ProSerLeuMetThrGlyArgSerAlaProSerProAlaLeuSerArgSerHisPro 1147
Db 183 CCTTCATGATGACTGGAAGATCTGCCCTCTCCCTCCCTGCTATCGAGGTCTCATCCT 242
QY 1148 ArgThrGlySerValGlnThrSerProSerSerThrProGlyThrGlyArgArgGlyArg 1167
Db 243 CGTACTGGGTCTCTCCAGACAGCCCATCAAGTACTCCAGTCCGACGACGAGGGGCCGA 302
QY 1168 GlnLeuProGlnLeuProProLysGlyThrLeuGluArgSerAlaMetAspIleGluGlu 1187
Db 303 CAGCTTCCACAGCTTCCACCAAGGGAACGTTGGATGAAAGCA----- 347
QY 1188 ArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGlySerAspProArgLeu 1207
Db 347 ----- 347
QY 1208 GluGlnAspTyrHisSerLysTyrArgSerGlyTyrAspProHisArgGlyAlaAspThr 1227
Db 347 ----- 347
QY 1228 ValSerThrLysSerSerAspSerAspValSerValSerAlaValSerArgThrSer 1247
Db 347 ----- 347
QY 1248 SerAlaSerArgPheSerSerThrSerTyrMetSerValGlnSerGluArgProArgGly 1267
Db 347 ----- 347
QY 1268 AsnArgLysIleSerValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGly 1287
Db 347 ----- 347
QY 1288 LysAsnLeuThrLysSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAsp 1307
Db 347 ----- 347
QY 1308 GlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgArgSer 1327
Db 347 ----- 347
QY 1328 SerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSer 1347
Db 347 ----- 347
QY 1348 GlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerThr 1367
Db 348 -----GGAGGTAAATAAACTAAGGAGCACTGTCCAAAGAGTACA 386
QY 1368 GluThrGlyLeuAlaVal-GluMetArgAsnTrp-----MetThrArgGlnAlaSerAr 1385
Db 387 GAAACAGGCGCTGCCCTGGGAAATGAGGAACCTGGKATGGACTCGACAGGCGAAGCGGAG 446
QY 1385 gGluSerThrAspGly-SerMetAsnSerTyrSerSer---GluGlyAsnLeuIlePheP 1404
Db 447 AGTCTTACAGATGGTTAGCATGAACAGCTAACAGCTCCAGAGGAANTCTGATTTTNC 506
QY 1404 ro 1404
Db 507 CN 508
```

RESULT 8

```
ABLI12714/c
ID ABL12714 standard; cDNA; 20604 BP.
XX
AC ABL12714;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 32624.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
```

```
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; AB868611.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 32624; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
XX sequences (ABLI01840-ABLI16175) and the encoded proteins
XX (ABBS7737-ABBY2072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 20604 BP; 5097 A; 4095 C; 5444 G; 5968 T; 0 other;

Alignment Scores:
Pred. No.: 3,55e-13 Length: 20604
Score: 470.50 Matches: 115
Percent Similarity: 57.42% Conservative: 32
Best Local Similarity: 44.92% Mismatches: 62
Query Match: 5.68% Indels: 48
DB: 23 Gaps: 5

US-09-617-099B-1 (1-1590) x ABLI12714 (1-20604)
QY 1373 ValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThrAspGlySerMet 1392
Db 1779 ATTGATGAGAGATGATACAAAGAAATATCCACACGACGCTGTACTAAAGTCTTCT 1720
QY 1393 AsnSerTyrSerSerGluGlyAsnLeuIlePheProGlyValArg-----Leu 1408
Db 1719 TTACTGTATCCCCACCCGACCCATTCAAATCCCTCATAGATTCCTCCCTTCGTTG 1660
QY 1409 AlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGly 1428
Db 1659 CGCAACGATGGAGCCCTCAATGATTCGTGGATGGCTGGCCCTCGTCACTGGTGGT 1600
QY 1429 ArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLys 1448
Db 1599 CGCCAGGTCTGGAGTCCCTCGCTGGCGGACATCCAGCTATCGTGTGCCACCAAAA 1540
QY 1449 GlyGlnLeuGluValGluIleIleArgAlaArgGlyLeuVal----- 1463
Db 1539 GGCTGTCTGGAGTGGAGGTCTATCGGGCCAGGGCTTACAGGTACGTAGCAGCATTCG 1480
QY 1464 -----LysProGlySer 1467
Db 1479 TATAGCATCTTATGATTTTCTAACACTCTATCTGTGATCCCGACCAAAAGCCCAA-TCA 1421
```


Db 1890 GTGACGAGCGGAGGAGTCCCACTCTAGATCCCGACCCACTAGGGGTAGATCT---CG 1946
Qy 355 uSerArgSpGluTyrGluArgGlnA:ArgGluGluGluTyrGlnAlaArgTyrArgSe 375
Db 1947 TTCTAGAACACACGCGCGCGGCGAGTCCCGCTCTAGAACACCTGCCAGCGGAGATC 2006
Qy 375 rAspProAsnLeuAlaArgTyrProValLysProGln----- 387
Db 2007 A-----CGATCCAGAACTCCACGAGCGTAGGTCTCGGTCTAGAACACCGAGCCGC 2057
Qy 388 -----ProTyrGluGluGlnMetArgGlnHisAlaGluVa 399
Db 2058 GAGGGGAGGTCTCGGTCTAGAACACCTGCTAGGCGCAGATCTAGACCCGATCACCAGT 2117
Qy 399 l-----SerArgAlaArgHisGluArgArgHisSerAspValSerLeu----- 413
Db 2118 ACGACGCGAGGTCTCGGTAGATCACCAGCCAGGAGAGTGGCAGGTCCAGCTCTAGAAC 2177
Qy 414 -----AlaAsnAlaGluLeuGluAspSe 421
Db 2178 CCCAGCTAGACGTGGCGCTCACGCTCCAGAACCCCGACGAGCGTGGCGCTCACGCTC 2237
Qy 421 rArgIleSerLeuLeuArgMetAspArg-----ProSerArg----- 433
Db 2238 TAGAACCCGACTAGACGCGAGTGGTCTCAGCTCCAGAACACCGACGCGAGGAGGAG 2297
Qy 434 -----GlnArgSerValSerGluArgAlaAlaMetGluAsnGl 447
Db 2298 GTCTCGGTCTAGGACACCAAGCAGGAGAGATCCCGCAGTAGAAGCTTAGTTAGACGTGG 2357
Qy 447 nArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnAr 467
Db 2358 AAGATCTCACTCT-----AGAACACCTCAAGAGAGGCGAGTCTGCTCATCTTCAGA 2411
Qy 467 gThrSerAsnHisSerProThrProArgSerProIleProLeuAspArgProAs 487
Db 2412 CGCGAATAACAAATCCAGAACATCTCAAGAGAGAGCAGG---TCCAATTCAGCCCGAGA 2468
Qy 487 pMetArgArgAlaAsp-----SerLeuArgLysGlnHisHisLeuAspProSerSerAlaVa 506
Db 2469 AATGAAGAAATCTCGCATTTCTTCAAGCGGAGCAGGTCTCTCTTCCACCGCTCCAA 2528
Qy 506 lArgLysThrLys-----ArgGl 512
Db 2529 AGCAAAATCTCGCTGTCTTTGAGCGCGAGCCTTTGAGGGTCTTCCCATCCCTAAGCA 2588
Qy 512 uLysMetGluThrMetLeuArgAsnAspSerLeuSerSerAspGln----- 527
Db 2589 AAAGTCACAGACACACCCAGCGCGAGTCGCTCTGGATCTCCCAACCTAAAGCTAAATC 2648
Qy 528 -----SerGluSerValArgProProProArg----- 537
Db 2649 TAGAACGCCACCCAGACGAGTCGCTCCAGTTCTTCTCCGCCACCTAAACAGAAATCTAA 2708
Qy 538 -----ProHis---LysSerLysLysGlyGl 545
Db 2709 GACACCATCAAGACAAAGTCATTCAGTTCATCTCTCATCTCTAAAGTAAATCTGGAAC 2768
Qy 545 yLysMetArgGlnValSerLeuSerSerGlu---GluGluLeuAlaSerThrProGl 564
Db 2769 ACCACCGAGGCGAAGGTCCTAACAAGTCCCGAGCCCAATGAGCAATCTGTAAACGCCACA 2828
Qy 564 uTyrThrSerCys-----AspAspValGluLeuGluSerGluSer----- 577
Db 2829 GAGACGGAGCTGTTTGAATCATCCTGACCTGAGTTGAAATCTAGGACCCCTTCTAG 2888
Qy 577 ----- 577
Db 2889 ACATAGCTGCTCAGGGTCTCTCTCTCTAGAGTGAAATCTAGCACACCTCCAGACAGAG 2948
Qy 578 -----ValSerGluLysGlyAs 583
Db 2949 CCCATCTAGGTATCATCTCCCAACCCCAAGTGAAGGCAATATATATCACCAGAGCAAG 3008

Qy 583 pSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSer----- 598
Db 3009 AAGCCATTCTGGCTCTCTTCTCAAGTCTCTAGTAGGGTACGTCGAGAACAACTCCACG 3068
Qy 598 ----- 598
Db 3069 GCGAAGCAGATCAGTATCTCTCCCTGCTCCAATGTGGAATCCAGATTGTGTCCAGATACAG 3128
Qy 599 -----AspSerAsnThrArgSerGluArgGlnLysLysArgMe 611
Db 3129 TCATTCTGGGTCTCTCTCACCAGATACCAAGTGAACCTGAAACACCGCCAAGACAAAG 3188
Qy 611 tTyrTyrGlyGlyHisSerLeuGluGluAspLeuGluTyrSerGluPro----- 627
Db 3189 TCATCAGGTCTATTTCACCATACCCAAAGTAAAGGCCCAAACTCCACCGGGGCCAAG 3248
Qy 628 -----GlnIleLysAspSerGlyValaAspTh 636
Db 3249 TCATTCTGGATCAAGTCACCATGTCCCCAAGAGAGTCTTAAGACTCAGTCTCAAG 3308
Qy 636 rCys-----SerSerThrThrLeuAsnGluGl 645
Db 3309 TTGCGCTGGATCCCTCTCTCTGTGCGAGGTAAATCTTAGCACACCACCGCGGAGAG 3368
Qy 645 uHis----- 646
Db 3369 CTATTTGGTGTCTCATCTCTGCAACTGAAGACAAATCTCAAACTTCAACGACCCACAG 3428
Qy 647 -----SerHisSerAspLys-His----- 652
Db 3429 ATCTGATCTTCAAGTCCAGAAAGTGAGACAGATCATTCAGATCACCATCTCTGCGAG 3488
Qy 653 -----ProValThr 656
Db 3489 CAAATCTCAAAACATCACCTAAGGAGGTCCGTCCAGTCTTTCATCTCCAGTCACTGAGCT 3548
Qy 656 rGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuL 676
Db 3549 GGCATCCAGATCTCCAATNAGACAAAGATAGA-GGTGAGTCTCTCAGCGAGTCTTATGTGA 3607
Qy 676 yAspGlySerValProArg----- 682
Db 3608 AATCTGGAATGCTCTCTGAGCAGCAGGAGTTCGAGTCTGACTCTTTCATATCTCTACAG 3667
Qy 683 --AspSerGlyAlaMetLeuGlyLeuLysValValGly-----G 695
Db 3668 TGGACTCGAATTTCTCTCTGGGGCAGAGTAG-ATTGGAGACTGCTGAATCAAAAGAGAAA 3726
Qy 695 lLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLys-ValLysLysGlySer 714
Db 3727 ATGGCTTACCCCTCAGGAGGA---TGCTACTGCATCACCTCTCTAGACAGAAACAAA 3783
Qy 715 LeuAlaAspThrValGlyHisLeuArgPro----- 724
Db 3784 TTTAGTCCCTTTCAGTACAGTAGGCTGCTTCACCTGCTTTCATGTTTCAAGACACACTT 3843
Qy 725 -----GlyAspGluValLeuGluTyrPasn 732
Db 3844 AGAACCCGCCAGAGAGAAAGTGTGTGCTGGTCACTCTCCAGAAACAAAGAGCAAAAT 3903
Qy 733 GlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGluSerLys 752
Db 3904 AGTGCATTGCTTACGTCAAGCCAGATGAAGAGTTAATGAGGTGGTAGAAGTCTGAA 3963
Qy 753 ProGluProGlnValGluLeuValSerArgProIleGlyAspIleProArgIlePro 772
Db 3964 -----GAACCCGAGGCCAAATC-----CTGCT 3987
Qy 773 AspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSer-----GlnLys 790
Db 3988 CATTTGCTTTCAGAACTTAAAGAAATGTCCACAGTAACTTTGAATCATCTCTCTGAAGTA 4047

ID AAL07156 standard; DNA; 24387 BP.
XX
AC AAL07156;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 9844.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX Homo sapiens.
XX WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227709.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251038.
PR 05-DEC-2000; 2000US-0251989.


```
Qy 581 sGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSer----- 598
Db 13435 ACAAGAAGCGCATTTCTGGCTCTCTCTCTCAAGTCTCTAGTAGGCTGAGTGGAGACAAC 13376
Qy 598 ----- 598
Db 13375 TCCAGCGGAGAGCATCAGTATCTCCCTGCTCCCAATGTGGATCCAGATTGTGGCAAG 13316
Qy 599 ----- 599
Db 13315 ATACAGTCATTTCTGGTCTCTCTCCAGATACCAAAGTGAACCTGAACACCCGCCAAG 13256
Qy 609 sArgMetTyrTyrGlyGlyHisSerLeuGluGluAspLeuGluTrpSerGluPro----- 627
Db 13255 ACAAGTCACTTCAGGGTCTATTTCCACATACCCCAAGTAAAGCCCAAACTCCACCGGG 13196
Qy 628 ----- 628
Db 13195 GCCAAGTCTTCTGGATCAAAAGTCACCATGTCCCAAGAGAGTCTAAAGACTCACTAGT 13136
Qy 634 lAspThrCys----- 643
Db 13135 TCAAAGTTGCCCTGAGTCCCTCTCTCTGTGCGAGGAGTAAATCTAGCACACCACCGG 13076
Qy 643 nGluGluHis----- 646
Db 13075 CGAGAGCTATTTTGGTGTCTCATCTCTGCAACTGAAGAGCAATCTCAAACCTTCCACAGA 13016
Qy 647 ----- 652
Db 13015 CCACAGATCTGATCTTCAAGTCCAGAAAGTGAGACAGAGTCAATTCAGAAATCACCATCTCT 12956
Qy 653 ----- 654
Db 12955 GCAGAGCAAACTCAAACATCACCTAAGGAGGTGGTCCAGGTCTTCTATCTCCAGTCCAC 12896
Qy 654 alThrTrpGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysA 674
Db 12895 TGAGTGCATCCAGATCTCCAATAAGACAAGATAGA-GGTGAGTCTTCAGCGAGTCTTA 12837
Qy 674 rGluLysAspGlySerValProArg----- 682
Db 12836 TGTGAAATCTGGAATGCTCTCTGAGCAGAGAGGTTCCAGTCTGCTCTTCTTCATATC 12777
Qy 683 ----- 694
Db 12776 CTACAGTGCAGTCCGAATCTCTCTTGGGCGAGTAG-ATTGGAGACTGCTGAATCAAAA 12718
Qy 695 ----- 712
Db 12717 GAGAAAATGGCCTTACCCCTCAGGAGGA---TGCTACTGCATCACTCTCTAGACAGAAA 12661
Qy 713 GlySerLeuAlaAspThrValGlyHisLeuArgPro----- 724
Db 12660 GACAAATTTAGTCCCTTCCAGTACAGGATAGGCTGAGTCTCACTGGTATTCAAAGAC 12601
Qy 725 ----- 730
Db 12600 ACACCTTAGAACCCCGCAAGAGAAAGAGTGGTGGTGCATCTCCAGAACAAAGAG 12541
Qy 731 TrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGlu 750
Db 12540 CAAATATAGTCATTCCTACGTCAAGCAAGATGAAGATTAAATGAGGTGTAGAGAG 12481
Qy 751 SerLysProGluProGlnValGluLeuValValSerArgProIleGlyAspIleProArg 770
Db 12480 TCTGAA----- 12457
Qy 771 IleProAspSerThrHisAlaGlnLeuGluSerSerSerPheGluSer----- 788
Db 12456 CTGTCTCTATTTGCTTCAGAACTTAAAGAAATGTCCACAGTAATCTTGAATCATCTCTCT 12397
Qy 789 GlnLysMetAspArgProSerIleSerValThrSerProMetSerProGlyMetLeuArg 808
Db 12396 GAAGTAGAAAGAGCTCTGTGTCTTACTCTTGATCAGAGCCAGTCACAGGCTTCT 12337
Qy 809 ----- 823
Db 12336 TTGGAAGCAGTAGAAGTCCCTTCAATGGCCTCATCT----- 12295
Qy 824 ----- 840
Db 12294 GGGCCACATTTTCTCCAGAACATAAAGAACTG----- 12256
Qy 841 LeuProSerArgGluAsp----- 855
Db 12255 TCCCACTCAGGAGAACAGCTTTGGATCACCTTTAGAAATTTAGAACTCAGGC----- 12202
Qy 856 IleTyrPheLeuProAspArgSerAspLysAsnLysArgArgThrIleValLysLys 875
Db 12201 ----- 12154
Qy 876 ThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPhe 895
Db 12153 GATTTG----- 12107
Qy 896 Arg----- 907
Db 12106 AGACATGAAGAACAATCGACAAGATCTCTTGACACAGCAGTCTCAGATTATCCCCAGA 12047
Qy 908 ----- 925
Db 12046 TGCAGTGGAAAAGCGAGGATGCTTCAATCAG----- 12008
Qy 925 uGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGlnThrHisAspValse 945
Db 12007 CTCTTCACTGTGTCTGATGCTATCCAGAACACCCCTCGAGAGAAAGAGTAGTCTTCTGC 11948
Qy 945 rSerLeuPro----- 959
Db 11947 ATCTTCTCTGAAATGAAGATGTTTACCAGAACTCCATCA----- 11900
Qy 959 gGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspse 979
Db 11899 AAGCAGGTCTGGTCTTCTCCAGACTTGA----- 11869
Qy 979 rGluValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAs 999
Db 11868 ----- 11840
Qy 999 nGlyArgAspLeuGlnSerSerThrLeuSerValProGlu---GlnValMetSerSerAs 1018
Db 11839 CCTGTCTGGGTCTCTCTGGAATGAAGATATACCTAGAACGCCATCTAGAGGGAGAG 11780
Qy 1018 nHisCysSerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTr 1038
Db 11779 CGAATGTGATTTCTCCCAAGAACCGAAGCTTTCCTCAGACTCTTAGCCGAGAGTCTG 11720
Qy 1038 pSerProSerAlaPro----- 1051
Db 11719 TTCTCCATCATCCCGAGAGCTCAACAACAAGTGTCTTACCCCGCAGAGAGAA----- 11668
Qy 1051 nGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnThrIle----- 1065
Db 11667 ----- 11612
Qy 1066 ----- 1075
Db 11611 GCAGAGAGTCTGTGGGATCTCTCAAGAACTTGTGTGAACCCAGTGCATCCCTCA 11552
Qy 1076 -AspHisTyrSerSerAspArgAspArgCysGluAlaAlaAspArgGlnProTyr-- 1094
Db 11551 GGAAAGAGTGTGAGTCTCTCTCCAGATTTCTTAAGCCCAAGACACGACCCACCTCTG 11492
Qy 1095 -HisArgSerArgSerThrGluGlnArgProLeuLeuGluArg----- 1111
```

```
Db 11491 GCAGAGGAGTCGGTCTGGATCATCTCCAGAGCTTCACAGCAAAATCTCGACTATCCCTCG 11432
Qy 1111 rArgSerArgSerGluArgProAspThr---AsnLeuMetArgSerMetProSerLe 1130
Db 11431 GCGCAGTAGGTCGTCTCCCTCGAGTGAAGTAAGCAAGCAGCAGCACCAGGGC 11372
Qy 1130 uMetThrGlyArgSerAlaProProSerPro-----AlaLeuSerArgSerHisProAr 1148
Db 11371 ACAGAGTGGTCTGTATCTCTCTCGAACCTTAAGCTCCAGCCCTCGGCCCTCCAG 11312
Qy 1148 g-----ThrGlyse 1151
Db 11311 ACAGAGCAGATCAGGTCCTCATCAAGCAAGGAGCAGAGCCCTCTCCTGAAGGAAGCAGCAG 11252
Qy 1151 rValGlnThrSerProSerSerThrPro-----GlyThrGlyArgArgGly---ArgG1 1168
Db 11251 TACGAGTCCTCTCTCGAACATCCGCCCAAAATCCAGAACTGCTCGCAGAGGTTCCAGGTC 11192
Qy 1168 nLeuProGln-----LeuProProLysGlyThrLeuGluArgSerAl 1182
Db 11191 ATCACCAGAGCCCAAGACCAAGTCTGTACACCACTCGA-----CGTCGAGGTC 11141
Qy 1182 aMetAspIleGluGluArgGlnMetLysLeuAsnLysTyLysGlnValAlaG1 1202
Db 11140 TCGATCATCTCGGAGCTAAACAAGGAGGCGCAGACTGTCCCGTAGAAGCGCTCTGCTC 11081
Qy 1202 ySerAspProArgLeuGluGlnAspTyHisSerLysTyArgSerGlyTrpAspPro-- 1221
Db 11080 ATCTCTCACCA-----GAAACTCGCTCTAGAACTCCCCCAAG 11045
Qy 1222 -HisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspValse 1241
Db 11044 GCACCGGAGAGTCCTCAGTCTCTCC---CCGAGGCGAGCGGCAAAATCGAGTCTTC 10988
Qy 1241 rAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyMetSer---Va 1260
Db 10987 ACGCCGAGCGGCTCAGCTTCATCTCCAGCACTAAGCAACCTCAAGGAGGCGGCTC 10928
Qy 1260 lGlnSerGluArgProArgGly-----AsnArgLysIleSerVa 1273
Db 10927 TCCTTCGCCAAAGCTCGTGGACTCCAGAGGTCCTCCCTCCCTCAAGGAGAGAGAAAAC 10868
Qy 1273 lPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSe 1293
Db 10867 AAGAACACCCGACGTCGAGATAGTCTGGATCTTCAGTCAACCTCTCGGCGAAGACA 10808
Qy 1293 rThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGly-----Se 1309
Db 10807 GCGGAGCCGGTCAAGGTCGGGTTACTCGCGGCGGAGGGAGGCTCTGGTTATCATCTC 10748
Qy 1309 rGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSerSerIl 1329
Db 10747 AAGGTCACTCGCCGCGAGGAAGTTCGCGACCTCTCTCGACGCCGAAGA----- 10696
Qy 1329 eGlyAlaLysMetValAlaIleValGlyLeuSerArg-----LysSerArgSerAl 1346
Db 10695 -----GGCGCTCTCGACACCCCAACCCGAGTCGGAACG 10661
Qy 1346 aSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSer----- 1361
Db 10660 TTCTCGCTCAGCAGCATCAGCAGCCCGCTGGAAAGCTCTAGATCTCGAGCTCTCCAGC 10601
Qy 1362 -ThrValGlnArgSer---ThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetTh 1380
Db 10600 CACTCAGCGGCGATCAGGTCAGACCAACCCCTG-----ATAAG 10562
Qy 1380 rArgGlnAlaSerArgGluSerThr 1388
Db 10561 CCGAGCTAGGTCAGATCTCGAACT 10537
RESULT 12
AAV90372/c
```

AAV90372 standard; cDNA; 355 BP.

AAV90372;

15-FEB-1999 (first entry)

EST clone DMI18.

Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
gene therapy; ss.

Homo sapiens.

WO9845436-A2.

15-OCT-1998.

10-APR-1998; 98WO-US06955.

10-APR-1997; 97US-0838821.

(GEMY) GENETICS INST INC.

Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
Racie LA, Spaulding V, Treacy M;

WPI; 1999-070077/06.

New polynucleotides encoding human secreted proteins - derived from
e.g. human blood, kidney, foetal lung, placenta, testes, brain,
ovary, pituitary, retina and colon cDNA libraries.

Claim 1; Page 519-520; 618pp; English.

The present sequence represents a human expressed sequence tag (EST).
The polynucleotide, which is a secreted EST, and the encoded protein
are predicted to have useful biological activities which would make
them suitable for treating, preventing or ameliorating medical
conditions in humans and animals, although no supporting data is
given. Suggested activities include nutritional activity, immune
stimulating or suppressing activity, haematopoiesis regulating
activity, tissue growth activity, activin/inhibin activity,
chemotactic/chemokinetic activity, haemostatic and thrombolytic
activity, receptor/ligand activity, anti-inflammatory activity,
cadherin/tumour invasion suppressor activity, tumour inhibition
activity. The polynucleotide may also be useful for gene therapy.

SQ Sequence 355 BP; 93 A; 82 C; 83 G; 97 T; 0 other;

Alignment Scores:

Pred. No.:	4.6e-10	Length:	355
Score:	368.00	Matches:	75
Percent Similarity:	96.34%	Conservative:	4
Best Local Similarity:	91.46%	Mismatches:	3
Query Match:	4.44%	Indels:	0
DB:	20	Gaps:	0

US-09-617-099B-1 (1-1590) x AAV90372 (1-355)

Qy 1272 SerValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThr 1291

Db 246 AGTGTCTTATCCAAATGCAAGCAGACAAATGGCATATCGGGAAGCATGCA 187

Qy 1292 LysSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSer 1311

Db 186 AAAAGCACCACATCATGTGGAGACATGTCTCATCTGGAGAGATGATGGCAGCGCT 127

Qy 1312 AspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSerSerIleGlyAla 1331

Db 126 GACACTGCGTGGGACCTTGGGACCACTGGCAGTGGCAAAAGGCGGCTCTAGCTTGGTGGC 67

QY 1332 LysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeuSerGln 1351
 Db 66 AAAATGGTAGCTATCGTGGTCTGTACGGAAAGTCGAGTCTTCTCAGTCACGGGA 7
 QY 1352 ThrGlu 1353
 Db 6 CTCGAG 1
 RESULT 13
 AAV89630
 ID AAV89630 standard; cDNA; 355 BP.
 AC AAV89630;
 XX
 DT 15-FEB-1999 (first entry)
 XX
 DE EST clone CO618.
 XX
 KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
 KW gene therapy; ss.
 OS Homo sapiens.
 XX
 XX WO9845436-A2.
 PN
 PD 15-OCT-1998.
 XX
 PF 10-APR-1998; 98WO-USO6955.
 XX
 PR 10-APR-1997; 97US-0838821.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie LA, Spaulding V, Treacy M;
 XX
 DR WPI; 1999-070077/06.
 XX
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries.
 XX
 PS Claim 1; Page 279; 618pp; English.
 XX
 CC The present sequence represents a human expressed sequence tag (EST).
 CC The polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.
 XX
 SQ Sequence 355 BP; 97 A; 83 C; 82 G; 93 T; 0 other;
 Alignment Scores:
 Pred. No.: 4.6e-10 Length: 355
 Score: 368.00 Matches: 75
 Percent Similarity: 96.34% Conservative: 4
 Best Local Similarity: 91.46% Mismatches: 3
 Query Match: 4.44% Indels: 0
 DB: 20 Gaps: 0
 US-09-617-099b-1 (1-1590) x AAV89630 (1-355)
 QY 1272 SerValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThr 1291
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 110 AGTGCTTTTACATCCAAAATGCAAGCAGACAAATGGGCATATCAGGAAGACATGACA 169
 QY 1292 LysSerThrSerIleSerGlyAspMetCysSerLeuGlyLysAsnAspGlySerGlnSer 1311
 Db 170 AAAAGCACCAGCATCAGTCGAGACATGTCTCACTCGAGAGAATATGCGACGCGCTT 229
 QY 1312 AspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSerIleGlyAla 1331
 Db 230 GACACTGCGAGTGGGCACCTTGGGCACCACTGGGCAAAAGCGCGCTCTAGCCTTGGTGCC 289
 QY 1332 LysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeuSerGln 1351
 Db 290 AAAATGGTAGCTATCGTGGTCTGTACGGAAAGTCGAGTCTTCTCAGTCACGGGA 349
 QY 1352 ThrGlu 1353
 Db 350 CTCGAG 355
 RESULT 14
 AAZ22301
 ID AAZ22301 standard; cDNA; 9551 BP.
 XX
 AC AAZ22301;
 XX
 DT 25-NOV-1999 (first entry)
 XX
 DE CDNA encoding a human trichohyalin (TRHY) protein.
 XX
 KW Human; trichohyalin; TRHY; protein; tissue structure; wound healing;
 KW terminally differentiating epidermal tissue; proteinaceous gel;
 KW breast implant; ss.
 XX
 OS Homo sapiens.
 XX
 PN US5958752-A.
 XX
 PD 28-SEP-1999.
 XX
 PF 14-FEB-1997; 97US-0800644.
 XX
 PR 30-APR-1993; 93US-0056200.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Kim I, Chung S, Park S, Steinert PM, Lee S;
 XX
 DR WPI; 1999-561041/47.
 DR P-PSDB; AAY30795.
 XX
 PT Human trichohyalin useful for forming a proteinaceous gel that promotes
 PT wound healing -
 XX
 PS Claim 1; Fig 3A-W; 126pp; English.
 XX
 CC The present sequence encodes a human trichohyalin (TRHY) protein.
 CC The protein is found in terminally differentiating epidermal tissue,
 CC and is involved in forming the structural architecture of such
 CC tissue. The trichohyalin protein is useful for forming a
 CC proteinaceous gel which may then be used for healing wounds, or in
 CC breast implants.
 XX
 SQ Sequence 9551 BP; 2907 A; 1974 C; 2870 G; 1800 T; 0 other;
 Alignment Scores:
 Pred. No.: 3.5e-08 Length: 9551
 Score: 360.50 Matches: 324
 Percent Similarity: 33.02% Conservative: 242
 Best Local Similarity: 18.90% Mismatches: 568
 Query Match: 4.35% Indels: 586
 DB: 20 Gaps: 63
 US-09-617-099b-1 (1-1590) x AAZ22301 (1-9551)

Db 4830 GGAGCCGACAGCGGGGAACAACGGTTTCTCCCGAGGAGGAGGAGAA-----GGAGCA 4883
Qy 722 LeuArgPro-----GlyAspGluValLeuGluTrpAsnGlyArg----- 734
Db 4884 GCGCGCGCGCCAGCGAGCGAGGAGGAGAGAGCTCCAGTCTCTGGAGGAAGAGGAGCA 4943
Qy 735 -----LeuLeuGlnGlyAla 739
Db 4944 GCTCCAGCGCGGAGCGTCCCAACAGCTCCAGGAGGAGGAGCGCTCCAGGAGGA 5003
Qy 740 ThrPheGluGluValTyraenilleLeuGluSerlyProGluProGlnValGluLeu 759
Db 5004 TCAGGAGAGGAGCGCAGCA-----GCAGCGCGCGACCAAAAATGGAGGTG 5051
Qy 760 ValValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeu 779
Db 5052 GCAACTAGAGA----- 5063
Qy 780 GluSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerVal-Th 799
Db 5064 -----AGAAAGGAGAGAGAGCGCGCCACACAGCTGTACGC 5096
Qy 799 rSerProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSe 819
Db 5097 CAAGCCAGCCCTACAAGAGAGCTGAGGAGAGAAACAGCAGCTGCTGCAGAGAGAG----- 5151
Qy 819 rIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLys 839
Db 5152 -----GAGGA 5156
Qy 839 sAspLeuProSerArgGluAspGlyA-ProArgAsnProTrpValLysIleTyPheLe 859
Db 5157 GGAGCTA----- 5163
Qy 859 uProAspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluPr 879
Db 5164 -----CAGAGAGAGAGCGCGAGAGAGAGAGCGCGCAAGAA----- 5199
Qy 879 oLysTrpAsnGlnThrPheIleTySerProValHisArgArgGluPheArgGluArgMe 899
Db 5200 -----CAGGAGAGACAATACCGCAG----- 5220
Qy 899 tLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGl 919
Db 5221 -----GAAGAGCAGCTGCAGCAGCAGGAGAGAGAGAGAGCTGCTGAG 5258
Qy 919 yGluIleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyK----- 937
Db 5259 AGAG-----GAACCGGAGAGAAAGAGAGCGCAGAGCGGAGAAAGCAATATCGGAA 5309
Qy 938 -----LysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerProTy 955
Db 5310 GGATAAGAGCTGCAGCAGAGAGAA----- 5334
Qy 955 rLeuProArgArgGlnLeuHisGlyGluSerPro-----ThrArgArgLeuGlnArgSe 973
Db 5335 -----GAGCAGCTGCTGGGAGAGCAACCGGAGAGAGAGAGAGCGCCAGAGCGGGA 5384
Qy 973 rLysArgIleSerAspSerGluValSerAspTyraAspCysGluAspGlyValGlyValva 993
Db 5385 GAAAAAATACCGCGAGGAGAG----- 5406
Qy 993 lSerAspTyraArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSerValProGluGl 1013
Db 5407 -----GAGTTGACGAGGAG----- 5429
Qy 1013 nValMetSerSerAsnHisCysSerProSerGlySerProHisArgValAspValIleGl 1033
Db 5430 GCTGCTG-----AGAGAGGAGAGCGGAGGA 5453
Qy 1033 yArgThrArgSerTrpSerProSerAlaProProProGlnArgAsnValGluGlnGlyHi 1053
Db 5454 GAGAAAGCGCCAGGA-GTGGGAGAGCGTACGTACCGCAAAAAGACGAGCTGCAGCAGGAG 5512

Qy 1053 sArgGlyThr-----ArgAlaThrGlyHisTyraenThr-IleSerArgMetAspArgH 1071
Db 5513 AAGAGCAGCTGCTGAGAGAGAGAACCGGAGAAAAGAGACTCCAGGAGCGGGAGAGCAAT 5572
Qy 1071 ieaArgValMetAspAspHisTySerSerAspArgAspArgCysGluAlaAlaAspA 1091
Db 5573 ATCGG-----GAGGAAGAGGAGCTGCAGCAGGAGGAAG 5605
Qy 1091 rGlnProTyraHisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThrT 1111
Db 5606 AGCAGCTGCTGGAGAGAGAACCGGAGAGAGAGCGCGCAGGAGCTGCAGAGCAATATC 5665
Qy 1111 hrArgSerArgSerSerGluArgProAspThrAsnLeuMetArgSerMetProSerLeuM 1131
Db 5666 GGAAGAGAGAGGAGCTGCAGCAGGAGAGAGAGCAGCTGCTGAGAGAGGA----- 5714
Qy 1131 eThrGlyArgSerAlaProProSerProAlaLeuSerArgSerHisProArgThrGlyS 1151
Db 5715 --ACCGAGAGAGAGAGCGCCGAGGAGCGGAGAGCAATGTCCGAGGAGAGGAGCTG 5772
Qy 1151 erValGlnThrSerProSerSerThrProGlyThrGly-ArgArgGlyArgGlnLeuPro 1170
Db 5773 CAGCAGGAGAGAGAGCAGCTGCTGCAGAGAGAGACCGGAGAGAGAGGCGCCAG----- 5826
Qy 1171 GlnLeuProProLysGlyThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArg 1190
Db 5827 -----GAGCTGGAGAGGAGCAATATCGGAGGAGAGAGAGCTTCAGCGC 5868
Qy 1191 GlnMetLysLeuAsnLysTyLysGlnValAlaGlySerAspProArgLeuGluGlnAsp 1210
Db 5869 CAGAAAGAGAAACAGCGATACCG-----GATGAGGAT 5901
Qy 1211 TyHisSerLysTyraArgSerGlyTrpAspProHis-ArgGlyAlaAspThrValSerTh 1230
Db 5902 CAGCGCAGTGATCTGAATGGCAGTGGACACAGAAAAGAAAATGCAGTTGCTGATAAC 5961
Qy 1230 r-----LysSerSerAspSerAspValSerAspValSerAlaValSerArg-- 1245
Db 5962 AAGGTTTACTGCAAGGCGAGAGAGAAATGAACAGTTCGCGCAGTTGGAGATTCGCCAGTG 6021
Qy 1246 ---ThrSerSerAlaSerArgPheSerSerThrSerTyMetSerValGlnSerGlu-- 1263
Db 6022 CGCGACAGACAATCCAGCAAGATCTGCAGCAGCTGCTGGGTGAACAGCAGAGAGAGAGAT 6081
Qy 1264 -----ArgProArgGly---AsnArgLysIleSerVa 1273
Db 6082 CGTGAGCAAGAGAGGAGCGCTGCAGCAGCAGCGCAACAGGCAATTTCCAGAGAGAAACAG 6141
Qy 1273 lPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSe 1293
Db 6142 CTGAGCGAGAGAGAGCAAGCAAGGAAGCAAAAGGCGCGACAGGAAGTCCCAAGAGAGAAAG 6201
Qy 1293 rThr----- 1295
Db 6202 CAGTTGCTGAGAGAGAGAAAGAGAGAGAGAGAGCGCGCTCAGAGACAGACAGAAATTC 6261
Qy 1295 rIleSerGlyAspMetCys----- 1301
Db 6262 CGCGAGGAGGAGACAGCTGCTCCAGGAAGGGAGGAACAGCGCTGCTCCCGCAGAGCGGT 6321
Qy 1301 ----- 1301
Db 6322 GACAGAAAATTCGCGAGAGAGAACTGCTCCATCAGGAACAAGGGAGAAATTCCTCGAG 6381
Qy 1302 -----SerLeuGluLysAsnAspGlySe 1309
Db 6382 GAGGAACAGCGGCTCGCGAGGAGAACGGGAGAGAAAATTCCTTAAGGAGGAGACAGCAGCTG 6441
Qy 1309 rGlnSerAspThrAlaValGlyAlaLeuGlyThr-----SerGlyLysLysArgAr 1326
Db 6442 CGCCTCGAGGAGCGGAGAGCACTCGTCCGAGGAGCGCGACAGAAAATTCGCGAGAGAGGA 6501

QY 1326 gSerSerIleGlyAlaIysMetValAlaIleValGlyLeuSerArgLysSerArgSerAl 1346
 Db 6502 CAGCAGCTGAGCGCCAGGCGGTGACAGAAATTCGTGAGAGGAAACAGCAGGTGGCC 6561
 QY 1346 aSerGln-----LeuSerGlnTh 1352
 Db 6562 CGCCAGGACGAGAGAGAAATTCCTGGAGGAGAACAGCAGCTGGCCGAGGAGCTCAC 6621
 QY 1352 rGluGlyGlyGlyLys----- 1358
 Db 6622 AGAAATTCGCGAAGAGGAGAACAGCTGCTCCAGAGAAAGGAGAACAGCAGCTGCACCC 6681
 QY 1359 -----LeuArgSerThrVa 1363
 Db 6682 CAGACGCTGACAGAAATTCCTGGAGGAGAACACAGCTGGCCGCGCAGAGCTGAC 6741
 QY 1363 lGlnArgSerThrGluThrGlyLeuAlaVal-----GluMetArgAsnTrpMetThrAr 1381
 Db 6742 AGAAATTCGCGAAGAGGAGAACCTGCGAGTACAGAACACAGAGAAATTCCTCGAGGAG 6801
 QY 1381 gGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnLe 1401
 Db 6802 GAACAGCAGCTGACCGCCAGCAGCAACGCGCAGAGAAATTCCTC-CAGGAGGAGACAGCAGCT 6860
 QY 1401 uillePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLe 1421
 Db 6861 GCGCGC-CCAGAGCGCGGCGAACACGCGGTCA----- 6893
 QY 1421 uGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly-AspIleG 1441
 Db 6894 -GGACCGTCACAGAAATTCGCGAGGAGGAGAACAGCTGCGCCAGGAGGAGGAGACAGC 6952
 QY 1441 lValGlyMetMet-----AspLysGlyGlnLeuGluValGluIleArgAlaA 1459
 Db 6953 AGCTGACGCGCCAGAGCGCTGACAGAAATTCGTTAGAGGAGAACAGAAAGTGGCGGCC 7012
 QY 1459 rGlyLeuValValLys-----ProGlySerLysT 1469
 Db 7013 A-GGAACAGAGAGAAATTCATGGAGGAGAACACAGCAGCTGGCCGCGCAGGAGGCCAA 7071
 QY 1469 hrLeuProAlaProTyrValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLysL 1489
 Db 7072 CAACAGCTGGCCA-----GGAGGACAGAAATTCGCGGA 7107
 QY 1489 yslYsThrLys-ValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSer--- 1507
 Db 7108 GACGACAGCTGCTCCAGGAGGAGGAGAACACAGCAGCTGCACCGCCAGAGCTGACAGA 7167
 QY 1508 -----PheGluGluSerProGlnGlyArg 1515
 Db 7168 AAATTCCTCGAGGAGAACCGCAGCTGGCG 7197
 RESULT 15
 ID AAV21511
 XX AAV21511 standard; DNA; 6755 BP.
 AC AAV21511;
 XX 17-AUG-1998 (first entry)
 DT DT
 XX Staphylococcal bacteriocin BacRI operon.
 DE DE
 XX BacRI; bacteriocin; antimicrobial; antibacterial; antibiotic;
 KW Moraxella bovis; infectious bovine keratoconjunctivitis; cancer;
 KW therapy; ds.
 XX Staphylococcus aureus strain UT0007 (ATCC 55800).
 OS OS
 XX W09812319-A1.
 PN PN
 XX 26-MAR-1998.
 PD PD
 XX 18-SEP-1997; 97WO-US16758.
 PF PF

XX 17-SEP-1997; 97US-0931999.
 PR 19-SEP-1996; 96US-0710561.
 XX (UNIV) UNIV KANSAS STATE RES FOUND.
 PA Crupper SS, Iandolo JJ;
 PI WPI; 1998-230316/20.
 DR Therapeutic proteinaceous substances from *Staphylococcus aureus* -
 PT useful to inhibit growth of wide range of prokaryotic or eukaryotic
 PT cells, e.g. *Moraxella bovis* causing infectious bovine
 PT keratoconjunctivitis
 XX Claim 2; Page 19-23; 38pp; English.
 FS This polynucleotide comprises the bacteriocin BacRI operon of
 CC *Staphylococcus aureus* UT0007. The sequence of the BacRI operon
 CC was determined by N-terminal sequencing of purified BacRI peptide
 CC (see AAW54171), with back-translation and plasmid analysis. The
 CC BacRI operon includes the BacRI gene (see AAV21510), a homologue of
 CC the *cylM* gene of the cytolyisin operon of *Enterococcus faecalis*
 CC whose function is involved in the maturation of pre-cytolysin,
 CC an ATP-transporter gene, *bio1* and *bio2* genes related to
 CC lactococcal biosynthesis and modification, and a gene involved in
 CC immunity function. BacRI peptides can be produced by construction
 CC of an expression vector containing an oligonucleotide or operon
 CC coding for BacRI, and use of the vector to transform host cells for
 CC BacRI expression. The entire BacRI operon has been cloned into
 CC plasmid pUB110, and *Bacillus subtilis* transformants secreted the
 CC recombinant BacRI peptide into the medium. Bacteriocin BacRI is
 CC active against many Gram-positive and Gram-negative organisms such
 CC as *Bordetella bronchiseptica*, *Pasteurella multocida* and
 CC *Staphylococcus aureus*; *Moraxella bovis*, causing infectious bovine
 CC keratoconjunctivitis, is especially sensitive. BacRI can also be
 CC used as an anti-cancer agent.
 XX Sequence 6755 BP; 3903 A; 1252 C; 1600 G; 0 U; 0 other;
 SQ

Alignment Scores:

Pred. No.: 6.18e-06 Length: 6755
 Score: 311.00 Matches: 288
 Percent Similarity: 34.2% Conservative: 226
 Best Local Similarity: 19.2% Mismatches: 593
 Query Match: 3.75% Indels: 399
 DB: 19 Gaps: 53

US-09-617-099B-1 (1-1590) x AAV21511 (1-6755)

QY 8 ArgGlyArgProAlaProThrProAlaAlaSerGlnProProGln-----ProGlu 25
 Db 61 AGAGAAAGAGAGCGGAGAACACACAGGGGGAGCGCCACGGGGAGGAGAACACCAAA 120
 QY 26 MetProAspLeuSerHisLeuThrGluGluGluArgLysIleIleLeuAlaValMetAsp 45
 Db 121 CAGCCCAACACCGGACAGCC-CGCGAGGAGAGGAGA-----AGC 158
 QY 46 ArgGlnLysLysGluGluLysGluGlnSerValLeuLysIleLys----- 61
 Db 159 AAAGACACACAGAGAGAGAACACCAACAAACAGAAACAGAAACCAACCAAGAGAA 218
 QY 62 -----GluGluHisLysAlaGlnPro 68
 Db 219 GAAACGACCCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 278
 QY 69 ThrGlnTrpPheProPheSerGlyIleThrGluLeuValAsnAsnValLeuGlnProGln 88
 Db 279 GCGAAGGGC-----GGAGACAAACGGGAAACAGCAA 308
 QY 89 GlnLysGlnProAsnGluLysGluProGlnThrLysLeuHisGlnGlnPheGluMetTyr 108
 Db 309 AAAAACAACAAACACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 368

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 08:13:46 ; Search time 12923 Seconds
(without alignments)
5033.375 Million cell updates/sec

Title: US-09-617-099B-1
Perfect score: 8285
Sequence: 1 MSAPLGRGRPAPTPAASQP.....TTRASQSLSSTGSPSYRS 1590

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2/1/USPTO.spool/US09617099/runat_14112003_191546_23681/app_query.fasta_1.1735
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -WAXLEN=200000000
-USER=US09617099@cgm_1_18872@runat_14112003_191546_23681 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htg.hum:*
40: em.htg.mus:*
41: em.htg.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8285	100.0	4980	6	AX137063 Sequence
2	8285	100.0	4980	6	B63816 Protein Rim
3	7930	95.7	4800	10	AB021131 Mus muscu
4	7459.5	90.0	5550	10	AF199331 Rattus no
5	7438.5	89.8	5592	10	AF199330 Rattus no
6	7420.5	89.6	5640	10	AF199322 Rattus no
7	6856.5	82.8	5172	10	AF199324 Rattus no
8	6543	79.0	4965	10	AF199329 Rattus no
9	6522	78.7	5031	10	AF199323 Rattus no
10	6522	78.7	5031	10	AF199326 Rattus no
11	6522	78.7	5031	10	AF199327 Rattus no
12	6522	78.7	5031	10	AF199332 Rattus no
13	6504	78.5	5079	10	AF199325 Rattus no
14	6261.5	75.6	5063	10	AF548738 Rattus no
15	6212	75.0	4851	10	AF199328 Rattus no
16	5638	68.1	3854	9	AB018294 Homo sapi
17	5239.5	63.2	4456	9	BC043144 Homo sapi
18	4583	55.3	5079	9	AY190519 Homo sapi
19	4513.5	54.5	5841	10	AF199333 Rattus no
20	4489.5	54.2	5655	10	AF007836 Rattus no
21	4402	53.1	5325	9	AB051866 Homo sapi
22	2936.5	35.4	6395	9	AB002338 Human mRN
23	2936.5	35.4	6638	6	BD083751 Nucleic a
24	2936.5	35.4	6638	6	BD097398 Nucleic a
25	2936.5	35.4	6638	9	AB045726 Homo sapi
26	2424.5	29.3	2223	9	AF263306 Homo sapi
27	2410.5	29.1	2301	9	AF263305 Homo sapi
28	2340.5	28.2	2103	9	AF263307 Homo sapi
29	1912	23.1	1877	9	AF007156 Homo sapi
30	1834.5	22.1	1791	9	AF263308 Homo sapi
31	1821	22.0	1704	9	AF263309 Homo sapi
32	1763.5	21.3	1551	9	AF263310 Homo sapi
33	1612	19.5	254783	2	AC101935 Mus muscu
34	1580	19.1	139527	2	AC141382 Rattus no
35	1498	18.1	139398	9	AC025836 Homo sapi
36	1498	18.1	152336	9	AF004714 Homo sapi
37	1498	18.1	162489	9	AF001572 Homo sapi
38	1452.5	17.5	4692	3	AF257062 Caenorhab
39	1221.5	14.7	1963	10	AF199335 Rattus no
40	1042	12.6	7239	9	D87074 Human mRNA
41	1042	12.6	7325	6	AX281663 Sequence
42	1025.5	12.4	1424	9	BC003103 Homo sapi
43	1022	12.3	1605	10	AF199334 Rattus no
44	830	10.0	1163	10	AJ310531 Mus muscu
45	828	10.0	11647	3	U41035 Caenorhabdi

ALIGNMENTS

RESULT 1

AX137063
 LOCUS AX137063 4980 bp DNA linear PAT 30-MAY-2001
 DEFINITION Sequence 2 from Patent EP1090986.
 ACCESSION AX137063
 VERSION AX137063.1 GI:14273409
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Seino, S., Shibasaki, T. and Ozaki, N.
 Protein rim 2
 Patent: EP 1090986-A 2 11-APR-2001;
 JCR PHARMACEUTICALS CO., LTD. (JP); Seino, Susumu (JP)
 FEATURES
 Location/Qualifiers
 1..4980
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 BASE COUNT 1503 a 1161 c 1282 g 1034 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,04e-252 Length: 4980
 Score: 8285.00 Matches: 1590
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-617-099B-1 (1-1590) x AX137063 (1-4980)
 QY 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 Db 32 ATGTCGGTCTCGCTCGGCGCCCGGGCGCGCGGCTCCACCCGCGCGCTCTCAACT 91
 QY 21 ProProGlnProGluMetProAsnLeuSerHisLeuThrGluGluArgGlyLeu 40
 Db 92 CCTCCGACCGCGAGATCGGACCTCAGCCACCTCAGCAAGAGAGAGGAAATCATC 151
 QY 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLys 60
 Db 152 CTGGCTGTCTCATGTCGTCTCAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGATC 211
 QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyLeuThrGlu 80
 Db 212 AAAGAAGAACACAAAGCAACACCGACACACAGTGGTTCCTTTAGTGGGATCACTGAACTG 271
 QY 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
 Db 272 GTAATAAAGTTCGTGAGCCCCAGCAAAACCAACCAATGAGAGAGAGAGAGAGAGAG 331
 QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGlnSerGln 120
 Db 332 CTGCACCAACATTTGAATGTTAAGAGGCAAGTCAAGAGATGGAGAGAGAGAGAGAG 391
 QY 121 GlnGlnGlnGluLysGlyAspAlaProThrCysGlyLysCysHisLysThrLysPhe 140
 Db 392 CAGCAGCAAG 451
 QY 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
 Db 452 GAGATGGATGGCGGCATTAATTTCTTATTCCTATTCCTAATTCCTAATTCCTAATTC 511
 QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
 Db 512 GTCGAGTGTCTTACCTCAACACAGAGTTATGTGGGTGTGAATTTGTGCCGAAACAA 571
 QY 181 GlnGluLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
 Db 572 CAAGAAATCCTCACTAAATCAGAGAGATGTTTATATATAGTGGGTCTTAACACACTGCAG 631

201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLysLys 220
 Db 632 CAACTGATCAAAAGGTTCTCGAGGCTTCGAATGAGGAGCCCTTCAGAGAGAGAA 691
 QY 221 AlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
 Db 692 GCAAAATCTACAGCAGCAGCCCACTTCCAAGGAGCCCACTTATCAGTACCTGCA 751
 QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrLeuLysAsnGlySerGly 260
 Db 752 GTTCAGAAAGCCAGCTCATGGCTCACAAGACAGGATCTATTAAATAATGGATCAGGA 811
 QY 261 VallyHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 Db 812 GTGAGCACCCAGATTGCCAGTCATGCTTCAGACAGAAACGAGTCCATCATGATGCC 871
 QY 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
 Db 872 AGGATCAAAATCGAAGTACGAGCAAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAT 931
 QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
 Db 932 CCTTCAGATGGTACATGCCAAGATCTCTTCGGATTTATGCTGATAGACGATCTCAGCGT 991
 QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
 Db 992 GAGCTCAATTTATGAGAACCTGGTCATTTAATTTACAGGATTTCTAACAGAGAGAGGC 1051
 QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
 Db 1052 CATAGACATTCCAAAGAGATATTTGTGATGATGAAGATGTGGAGAGCAGAGATGAATAT 1111
 QY 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
 Db 1112 GAAAGACAAAGAGAGAGAGAGAGATACAGGACGCTACAGAGATGATTCCTCAATCTGCGC 1171
 QY 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
 Db 1172 CGGTATCCGTAAAGCCACAACTTACGAGACAAATGCGCATCCACGCTGAGGTGTCC 1231
 QY 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
 Db 1232 AGGCAAGCAGATGAGAGAGAGCAGATGATGTTCTTTGGCAACGCTGACTAGAGAT 1291
 QY 421 SerArgIleSerLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
 Db 1292 TCCAGGATTTCTCTGCTAAGGATGGATAGACCATCAGGCAAGATCTGTATCTGAACGT 1351
 QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
 Db 1352 AGAGCTGCATGGAAACCAACGATCGTATTCAATGGAAAGACTTCGAGAGGCTCAGGGA 1411
 QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
 Db 1412 CAAAGTCTTATCCAAAGGACCTCAATCATAGTCTCCACCCCTCGCGGAGCCCT 1471
 QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
 Db 1472 ATACGCTTGTATAGACACAGATGAGCGCGCTGACTCTCCCTACGGAACAGCACCACTTA 1531
 QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
 Db 1532 GATCCAGCTCTGCTGTGGAGAAACGAGGAGAAATAATGGAAACCATCTTTAAGGAAT 1591
 QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
 Db 1592 GATCTTTGAGTTCAGACCACTCCGAGTCAGTGGCGCGCCGCCACCAAGGCTCATATA 1651
 QY 541 SerLysLysGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAla 560
 Db 1652 TCCAGAAAGAGAGGTAAATGCGCAGGTTTCACTGAGCAGCTCGAGAGAGAGCTGCA 1711
 QY 561 SerThrProGluTyrThrSerCysAspValGluLeuGluSerGluSerValSerGlu 580

Db 1712 TCCACACCTGAGTATCAAGCTGTGATGATGTCGAGCTGGAAAGCGAGAGTGTGAGTGAG 1771
Qy 581 LysGlyAspSerGlnLysGlyArgLysThrSerGlnGlnGlyValLeuSerAspSer 600
Db 1772 AAAGGGGACAGCTCAAAAGGGGAAAAGAAAAGTACTAGTGAGCAGGAGGATTTGTGCGATCT 1831
Qy 601 AsnThrArgSerGluArgGlnLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1832 AACACCGAGTCTGAGAGACAAAGAAAGAGATGTACTATGTTGGTGGCCACTTTTGGAGAG 1891
Qy 621 AspLeuGlnTrpSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 640
Db 1892 GATTGGAAATGGTCTGAGCCTCAGATTAAAGACTCTGGGGTAGATACCTGTAGTAGCACA 1951
Qy 641 ThrLeuAsnGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1952 ACCCTTAAACGAGGAGCATAGCCATAGTAGTAAGACACCCCTGTGACCTGGCAGCCATCCAAA 2011
Qy 661 AspGlyAspArgLeuLeuGlyArgGlnLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 2012 GATGGAGATGCCCTAATTTGGTTCGTATTTTAAATTAAGCGTTTAAAGATGGAGATGTA 2071
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 2072 CCTCGAGACTCAGGAGCAATGCTGGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA 2131
Qy 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 2132 GGTGCGATTTGTGCAATTTATTAACCAAGTAAAAAAGGAAGTTTGTAGCTACTGTAGGA 2191
Qy 721 HisLeuArgProGlyAspGluValLeuGlnTrpAsnGlyArgLeuLeuGlnGlyValThr 740
Db 2192 CATCTTAGACCGAGTGATGAAGTCTTGAATGGAAATGGAGCGCTATTTCAGAGGAGCCACA 2251
Qy 741 PheGluGluValTyrAsnIleLeuGluSerLysProGluProGlnValGluLeuVal 760
Db 2252 TTTGAGGAAGTTTACAACTATTATCTAGAATCCAAACCTGGAACCAAGTTGAGCTTGT 2311
Qy 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2312 GTTTCAAGGGCAATTTGGAGATATTTCTAGATATCTGATAGCAGCATGCAACAACCTGAAA 2371
Qy 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2372 TCCAGTTCTAGCTCATTTGAATCTCAAAAATGACCGCTCTCTATATCGTTTACCTCA 2431
Qy 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2432 CCCATGAGTCTGCGCATGCTGAGGGATGTCCTCGCAGTCTTATCTGGACAGCTTCAATA 2491
Qy 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2492 AAACATATGGTTTGAACAAGTTGGTTCACCAAGTTGATAGTTACAAATTTTGGGAGCAAGGAT 2551
Qy 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2552 CTCCCTTCCAGGGAAGATGGAGGCCCAAGGAATCTTATGTTAAGATTATCTTCCCTCCA 2611
Qy 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2612 GATAGAGTGATAAAAATAGAGAGAAACAAAACACAGTCAAGAAAATCTTTGGAACCCAAA 2671
Qy 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2672 TGGAAACAGACATTTCAATTTATCTCTCTGTCACCGAAGAGAATTCCTGTAACGAATGCTG 2731
Qy 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGlnGluSerGluPheLeuGlyGlu 920
Db 2732 GAAATTAACCTTTTGGGATCAAGCTAGAGTTTCGAGAAGAAGAGAGCGAATTTCTTAGGAGAG 2791
Qy 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspGluProHisTyrTrpTyrLysLeuGln 940

Db 2792 ATTTTAATTGAATTGGAAACAGACTTTTCTAGATGATGAGCCGCACTGGTATTAAGCTCAG 2851
Qy 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
Db 2852 ACCCATGATGTCTCTCAATTCGCCACTCCCTGCCCTTCCCATATCTTGCCTCCGAGCGAG 2911
Qy 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db 2912 CTCATGGAGAGAGCCCAACGCCAGCGCTGCAAAAGTTCGAAGAGATTAAGTGCAGTGAA 2971
Qy 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
Db 2972 GTGTCTGACTACGACTCGCAGCATGGCTGGAGTAGTGTGCAGATTATCGACACAATGGC 3031
Qy 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
Db 3032 CGCGATCTTCAAAAGCTCCAGCTTGTGCGTCCAGAACAACTCATGTCTCAAAATCATATGC 3091
Qy 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
Db 3092 TCACCATCAGGTTCTCTCATCGAGTAGATGTTATAGGAGGACAGGTCAATGTCGCT 3151
Qy 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 3152 AGTGGCCCTCTCTCAAGGAATGTGGAACAGGGGCACCGAGGACACGCTGCTACTGGC 3211
Qy 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 3212 CATTAACAACAATTAAGCCGATGGATAGACACCGTGTCTATGGATGACCACTACTCTTCA 3271
Qy 1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
Db 3272 GATAGAGACAGGATTTGTGAAGCAGCAGATGACAGCCATATCACAGTCCAGATCAACA 3331
Qy 1101 GluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAsp 1120
Db 3332 GAACAACGGCTCTCTCTAGAGCGGACCAACCCCGCTCCAGATCTCTTGAACGCTCTGAT 3391
Qy 1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
Db 3392 ACAAACTCATAGAGTCGATGCTCTTCAATTAAGTCTGGAAGATCTGCCCTCTCTTCACT 3451
Qy 1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
Db 3452 GCCTTATCGAGGTCTACCTCTGTCGTCGAGGTCTGTCCAGACAAGCCCATCAAGTACTCCG 3511
Qy 1161 GlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProLysGlyThrLeuGluArg 1180
Db 3512 GGAACAGGACGAAGGGCCCGCAGCTTCCACAGCTTCCACCAAGGGAAACATTGGAGAGA 3571
Qy 1181 SerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
Db 3572 AGTGCTATGGATATAGAGGAGAGAAATCGCAAAATGAACTTAACAAATACAAACAGTA 3631
Qy 1201 AlaGlySerAspProArgLeuGlnAspTyrHisSerLysTyrArgSerGlyTrpAsp 1220
Db 3632 GCGGATCAGACCCGACACTGGAGCAGATTTACATTCGAAGTATCTCCTCAGGATGGAT 3691
Qy 1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspVal 1240
Db 3692 CCACATAGAGGGCGACATATCTGTTTCCACTAAATCTCTCGCAGCATGATGTAAGTATGTA 3751
Qy 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerVal 1260
Db 3752 TCTCGGTTTTCAAGGACTAGTAGTGTCTTCGTTTCAGCAGCACAAGTACATGTCGCTC 3811
Qy 1261 GlnSerGluArgProArgGlyAsnArgIleSerValPheThrSerLysMetGlnAsn 1280
Db 3812 CAATCAGAGCGCGAGAGGAAACAGGAAATCAGTGTCTTTTATCTTCAAAATTCGAAAC 3871
Qy 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMet 1300
Db 3872 AGACAGATGGGCGTGTTCGGGGAAGAACTTGAACCAAAAGCACAGCATCAGTGGAGACATG 3931

QY	1301	CysSerLeuGluValAsnAspGlySerGlnSerAspThrAlaValAlaLeuGlyThr	1320
Db	3932	TGCTCAGTGGAGAGATGACCGCAGCCAGTCCGACACTGCAGTGGCGCCCTGGGTACC	3991
QY	1321	SerGlyLysLysArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSer	1340
Db	3992	AGTGCAAGACGCGGATCTAGCATTTGGGCGCAAAATGGTAGCTATTGTTGGTCTCTCA	4051
QY	1341	ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyGlyLysLysLeuArg	1360
Db	4052	CGGAAAGTCGAGTGCCTCTCAACTCAGCAAAACCGAAGGAGGAGGTAAGGATACGG	4111
QY	1361	SerThrValGlnArgSerThrGlnThrGlyLeuAlaValGluMetArgAsnTrpMetThr	1380
Db	4112	AGCATGTTTCAGAGAAGCCGAGACCGGCTAGCAGTGGAGTTCAGGAACCTGGATGACC	4171
QY	1381	ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn	1400
Db	4172	CGCCAGGCCAGCGGGAATCCACAGATGGCAGATGAACAGCTATAGCTCGAAGGAAT	4231
QY	1401	LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly	1420
Db	4232	CTGATCTTCCCTGGGCTCGCCTCTGACAGCCAGTTCAGTGAATTCCTGGATGCG	4291
QY	1421	LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetClyAspIle	1440
Db	4292	CTGGGCCCTTGCTCAGCTAGTGGAGCGCCAGACCTCTGGCTACTCTCGCAATGGGTGACATT	4351
QY	1441	GlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArgGly	1460
Db	4352	CAGGTGGGAATGATGATAAAAGGACAGCTGGAGGTAGAATCATCCGGCGCGCGGC	4411
QY	1461	LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu	1480
Db	4412	CTTGTGGTAAACCCAGGTTCCAGACACTGCCAGCACCGTATGTCAAGGTGTATCTGTTA	4471
QY	1481	AspAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeuGluPro	1500
Db	4472	GCACACGAGTCTGATACCCAAAAGAAAACCAAGGTGGCGAGAAAGCCCTGGAGGCC	4531
QY	1501	LeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleIle	1520
Db	4532	CTGTACACGACGCTCTTGCCTTCAGGAGAGACCCCGAGGAGGGGTGTACAGATCAT	4591
QY	1521	ValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetClyValAlaGlnIleLeu	1540
Db	4592	GTCTGGGAGATTATGGTGGTATGGATCAAAATCCTTTATGGGAGTGGCCAGATATCTC	4651
QY	1541	LeuAspGluLeuGluLeuSerSerMetValIleGlyTrpPheLysLeuPheProSer	1560
Db	4652	TTAGATGAACCTGAACCTATCCAACTGGGTGGATGGTGGTGGTGGTGGTGGTGGTGGT	4711
QY	1561	SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGlu	1580
Db	4712	TCCCTAGTAGATCCAACTCGGACCTCTGACACAGAGAGCTTCCCAATCGTCTCTGAA	4771
QY	1581	SerSerThrGlyProSerTyrSerArgSer	1590
Db	4772	AGTTCTACGGACCTTCTACTCTCGTTCA	4801
RESULT 2			
E63816			
LOCUS	E63816	4980 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Protein Rim2.		
ACCESSION	E63816		
VERSION	E63816.1 GI:22553654		
KEYWORDS	JP 2001103974-A/1.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 4980)		
	Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus.		

AUTHORS	Kiyono, S., Shibazaki, T. and Ozaki, N.		
TITLE	Protein Rim2		
JOURNAL	Patent: JP 2001103974-A 1 17-APR-2001; SUSUMU KIYONO, NIHON CHEMICAL RESEARCH K K		
COMMENT	OS Mus musculus (mouse)		
	PN JP 2001103974-A/1		
	PD 17-APR-2001		
	PF 08-OCT-1999 JP 1999288372		
FEATURES	PI SUSUMU KIYONO, TADAO SHIBAZAKI, NOBUAKI OZAKI		
	PC C12N15/09, C07K14/47, C07K16/18, C12Q1/68, G01N33/53//C12P21/08, C12N15/00		
	CC		
	FH		
source	Key	Location/Qualifiers	Location/Qualifiers
BASE COUNT	1503 a	1161 c	1282 g 1034 t
ORIGIN			
Alignment Scores:			
Pred. No.:	3,04e-252	Length:	4980
Score:	8285.00	Matches:	1590
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-617-099B-1 (1-1590) x E63816 (1-4980)			
QY	1	MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro	20
Db	32	ATGTCGGCTCCGCTCGGGCCCGGGCGCGCCGCTCCACCCCGGGCGCTCTCAACCT	91
QY	21	ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIleIle	40
Db	92	CCTCCGAGCCCGAGATGCCGACCTCAGCCACTCTCAGGAGAGAGAGAGAAATCATC	151
QY	41	LeuAlaValMetAspArgGlnLysLysGluGluGluGluGluGluGluGluGluGluGlu	60
Db	152	CTGGCTGTCTATGGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	211
QY	61	LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu	80
Db	212	AAAGAGAAACAAAGACCAACCCGACACAGTGGTTCCTTTAGTGGGATCACTGAACTG	271
QY	81	ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys	100
Db	272	GTAATAACGTTCTGAGCCCGCCAGCAAAACACCCATAGAGAGAGAGAGAGAGAGAG	331
QY	101	LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln	120
Db	332	CTGCACCAACAATTTGAAATGTAAGAGAGCAAGTCAAGAGAGAGAGAGAGAGAGAG	391
QY	121	GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe	140
Db	392	CAGCAGCAAGAGCAGAGAGGTGATGCCCGACCTGTGGCATCTGCCACAGAGCAAAATTT	451
QY	141	AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly	160
Db	452	CGAGATGATCGGCCATATTTGTTCTTATTCCTATTCCTATTCCTATTCCTATTCCTATTC	511
QY	161	GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln	180
Db	512	GGTCGAGTGCTTTACGCTCAACAGAGTATGTTGGGTGTGTATTTGTCGCGGAAACAA	571
QY	181	GlnGlnLeuLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln	200
Db	572	CAAGAAATCTCTCACTAAATCAGGAGCATGGTTTATATATAGTGGGTCTAACACACTG	631
QY	201	GlnProAspGlnLysValProArgGlyLeuArgGlnGluAlaProGlnGluLysLys	220

Db 632 CAACCTGATCAAAAGGTTCTCGAGGGCTTCGAAATGAGGAAGCCCTCAGGAGAGAA 691
Qy 221 AlalysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
Db 692 GCAAACTACACGAGCAGCCCGAGTTCACAGAGAGCCCGAGTGACTTATCAGTACCTGCA 751
Qy 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
Db 752 GTTGAGAAAGCCGAGCTCATGGGCTCACAAGACAGGATACTATTAAAAATGGATCAGGA 811
Qy 261 VallysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
Db 812 GTGAAGCACCCAGATGGCCAGTGACATGCCCTTCAGACAGAAAACCGAAGTCCATCAGTGTC 871
Qy 281 ArgAspGlnAanArgArgTyrGluGlnSerGluArgGluAspTyrSerGlnTyrVal 300
Db 872 AGGATCAAATCGAAGTACAGCAAGTGAAGAGAGAGGACTCTCAGTATGTT 931
Qy 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
Db 932 CCTTCAGATGTACAATGCCAAGATCTCTTCGGATTATGCTGATAGACGATCTCAGCGT 991
Qy 321 GluProGlnPheTyrGluGlnProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
Db 992 GAGCCTCAATTTATGAAGAACCTGGTCAATTTAAATTAACGGGATTTCTAAACAGGAGAGGC 1051
Qy 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgArgAspGluTyr 360
Db 1052 CATGACATTCCAAAGAGTATATTGTGGATGATGAGATGTGGAGAGCAGATGAATAT 1111
Qy 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
Db 1112 GAAAGACAAAGAGAGAGAGGAGAAATACCAGGACCGCTACAGAGTGTATCCAAATCTGGCC 1171
Qy 381 ArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSer 400
Db 1172 CGGTATCCGTAAAGGCCAACCCCTACGAAGACAAATGCGCATCCACGCTGAGGTGTC 1231
Qy 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuAsp 420
Db 1232 AGGCACGACATGAGAGAGGACAGTATGTTCTTTGGCAACGCTGAACTGAAAGAT 1291
Qy 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
Db 1292 TCCAGGATTTCTCTGTAAGGATGATAGACCATCAGGCAAGATCTGATCTGAAGCT 1351
Qy 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
Db 1352 AGAGCTGCATGGAAAAACCAACGATCGTATTCAATGGAAAGAACTCCGAGAGGCTCAGGGA 1411
Qy 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
Db 1412 CAAAGTCTTATCACAAAGAGACCTCAAAATCATAGTCTCCACCCCTCGCGGAGGCCCT 1471
Qy 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
Db 1472 ATACCGCTTGATACACAGACATGAGCGCGCTGACTCCCTACGGAACACACCACTTA 1531
Qy 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
Db 1532 GATCCCGAGCTGCTGTGAGGAAAAACCAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1591
Qy 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
Db 1592 GATCTTTGAGTTCAGACCATGCGAGTCAAGTGGAGGCGGCCGCCACCAAGGCCCTCAAAA 1651
Qy 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560
Db 1652 TCCAGAAAGAGAGTAAATGCGCAGGTTTCACTGAGCAGCTCGAGAGAGAGCTGGCA 1711
Qy 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
Db 1712 TCCACACCTGAGTATACAGCTGTGATGATGTGGAGCTGGAAAGCGAGAGTGTGAGTGAG 1771

Qy 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1772 AAAGGGACAGTCAAAGGGGAAAGAAAGAACTAGTGAGCAGGAGTTTTGTGCGATTCT 1831
Qy 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1832 AACACCAAGTCTGAGACACAAAGAAAAGGATGTACTATGTTGGCCACTCTTTGGAGAG 1891
Qy 621 AspLeuGlnTrpSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
Db 1892 GATTTGGATGGTCTGAGCCTCAGATTAAAGACTCTCGGGTAGATACCTGTAGTAGCACA 1951
Qy 641 ThrLeuAsnGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1952 ACCCTTAAACGAGGAGCATAGCCATAGTAGTAAGCACCTGTGACCTGGCAGCCATCCAAA 2011
Qy 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 2012 GATGAGATCGCTAAATTTGGTCTGATTTTATTAATAAGCGTTTAAAGATGGAGTGTA 2071
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 2072 CCTCGAGACTCAGGAGCAATGCTGGGCTTAAAGGTTCTAGGAGGAAGATGACTGAATCA 2131
Qy 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 2132 GGTGCACTTTGTGATTTTATCCAAAGTAAAGAAAGAGTTTAGCTGATCTGTAGGA 2191
Qy 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 2192 CATCTTAGACAGGTGATGAGTCTTGGATGGAATGGAGGCTATTGCAAGGAGCCACA 2251
Qy 741 PheGluGluValTyrAsnIleLeuLeuGluSerLysProGlnProGlnValGluLeuVal 760
Db 2252 TTTGAGGAGTTTACAACTATTCTAGAACTCCAAACCTCAACCAAGTTGAGCTTGT 2311
Qy 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2312 GTTTCAGGCGCAATTGGAGATATCTCTAGATATCTCTAGCAGCATGCAACTCGAA 2371
Qy 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2372 TCCAGTTCTAGCTCATTTGATCTCAAAAATGGACCGTCTCTCTATATTCGTTACTCA 2431
Qy 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2432 CCCATGAGTCTCGCATGCTGAGGAGATGTCGCGAGTTCTTATCTGACAGCTTCAATA 2491
Qy 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2492 AAATATGGTTTGACAAAGTTGGTCCACAGTTGATAGTTACAATTTTGGGAGCAAGGAT 2551
Qy 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2552 CTCCCTTCAGGAGAGATGGAGGCCAGGAGATCTTATGTTAGATTACTTCTCTCCA 2611
Qy 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2612 GATAGAGTGTATAAAAATAAGAGAGAAACAAAACACTCAAGAAAACCTTTGGACCCAAA 2671
Qy 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2672 TGGACACAGACTTTCATTATTCTCTGTCACCGAAGAGATTCCTGTGACAGGATCTG 2731
Qy 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920
Db 2732 GAATATCCCTTTGGGATCAAGCTAGAGTTCGAGAGAGAGAGAGGAGGAAATTTTAGGAGAG 2791
Qy 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisIleTyrLysLeuGln 940
Db 2792 ATTTTAATTCGAATGGAAACAGCTTTGCTAGATGATGAGCGCCACTGGTATATAGCTGCAG 2851

941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgGln 960
2852 ACCCATGATGTCCTCATTTGCCACTTCCCTCGCCCTTCCCATATATGCCCCGGAGGCAG 2911
961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
2912 CTCATGGAGAGCCCAACGCGCAGCTGCAAGGTCCAGAGAAATAGTGCAGTGAA 2971
981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisGlnGly 1000
2972 GTGTCTGACTACGACTGCGAGGATGCGTGGGAGTAGTGTGAGATTATGCACAAATGCG 3031
1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerHisCys 1020
3032 CGCGATCTTTCAAAAGCTCCACGTTGTCGTGCGCCAGAACAAAGTCATGTCAAAATCATTCG 3091
1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTyrSerPro 1040
3092 TCACCATCAGGGTCTCTCATCGAGTAGATGTTATAGGAGGACAGGTCACTGCTCCCT 3151
1041 SerAlaProProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
3152 AGTGCCCTCTCTCTCAAGGAATGTGAACAGGGCCACCGAGGGACAGTGTACTGCGC 3211
1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
3212 CATTAACAACACATTAAGCGAATGGATAGACACCGGTGTATGGATGACCACTACTCTTCA 3271
1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
3272 GATAGAGACGGATTGTGAGCAGCAGATAGACAGCCATATCACAGATCCAGATCCACA 3331
1101 GluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAsp 1120
3332 GAACAAGCGCTCTCTAGAGCGACCAACCCGCTCCAGATCCTCTGAAGCTCTGAT 3391
1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
3392 ACAAACTCATGAGGTGCGATGCTTCAATTAATGACTGGAGATCTGCCCTCTCTCACCT 3451
1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerThrPro 1160
3452 GCCTTATCAGGFTCTCACCTCTGACCGGTCTGTCCAGACAGCCCATCAAGTACTCCG 3511
1161 GlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
3512 GGBACAGAGCAGAGGCGGCGACAGCTTCCACAGCTTCCACCAAGGGAACATTTGGAGAGA 3571
1181 SerAlaMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
3572 AGTGCTATGATATAGAGAGAGAAATCGCAAAATGAAACTTAACAAATACAAACAGGTA 3631
1201 AlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTyrAsp 1220
3632 CGCGATTCAGACCCAGACTGGAGCAGATTAACATTCGAAGTATCGCTCAGATGGGAT 3691
1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspVal 1240
3692 CCACATAGAGGGGCAGACTACTGTTCCATAATCTCCGACAGTGTAGTAACTGATGTA 3751
1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerVal 1260
3752 TCTCGGTTTCAAGGACTAGTAGTCTCTCGTTTCAGCAGCAGCAAGTACATGTCCTGTC 3811
1261 GlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsn 1280
3812 CAATCAGAGCGCGCAGAGAGAAACAGGAAATCAGTGTCTTTTACATCCAAATGCAAAAC 3871
1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMet 1300
3872 AGACAGATGGCGTGTGGGGAAGAACTTACCAGAGCCAGCATCAGTGAGACATG 3931
1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320

3932 TGCTCATCTGAGAGAGAAATGACGCGACCGAGTCCGACACTGCAGTGGGCGCCCTGGGTACC 3991
1321 SerGlyLysLysArgArgSerSerIleGlyValLysMetValAlaIleValGlyLeuSer 1340
3992 AGTGGCAAGAGCGCGCATCTAGCAATTGGGCGCAAAATGGTAGCTATTGTTGGTCTCTCA 4051
1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyGlyLysLysLeuArg 1360
4052 CGGAAAGTTCGAGTGCCTCTCAACTCAGCCAAACCGAAGGAGGAGGTAAAGAGCTACGG 4111
1361 SerThrValGlnArgSerThrGluThrGluLeuAlaValGluMetArgAsnTyrMetThr 1380
4112 AGCATCTGTTTCAGAGAAGCAGGAGACCGGGCTAGCAGTGGAGATCAGGAACCTGGATGACC 4171
1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn 1400
4172 CGCAGGCGCAGCGCGGAATCCACAGATGGCAGCATGAACAGCTATAGCTCGAAGGAAT 4231
1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
4232 CTGATCTTCCCTGGGCTCGCTCTGACAGCCAGTTCAGTCAATTTCTCGATGCGC 4291
1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
4292 CTGGGCGCTGCTCAGCTAGTGGGACCGCAGACCCCTGGCTACTCTGCAATGGGTGACATT 4351
1441 GlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArgGly 1460
4352 CAGTGGGGAATGATGATAAAAGGAGCAGCTGGAGGTAGAATCATCCGGGCGCGCGC 4411
1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu 1480
4412 CTGTGTGTAACCAAGGTTCCAAAGACACTGCCAGCAGCTATGTCAAGGTATCTGTTA 4471
1481 AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeuGluPro 1500
4472 GACAAAGGAGTCTGATAGCCAAAGAAACCAAGGTGGCGAGAAAGCCCTGGAGCCC 4531
1501 LeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleIle 1520
4532 CTGTACAGCAGCTCTGTCTCTTCAGGAGAGCCCGCAGGAGGGGTGTACAGATCAT 4591
1521 ValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
4592 GTCTGGGAGATTATGTTGTTGATGATCACAATCTTTATGGAGTGGCCCGCAGATCTC 4651
1541 LeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPheProProSer 1560
4652 TTAGATGAATCGGAATATCCAACTATGGTATGGATGGATTCCTTCCCTCTCTTCC 4711
1561 SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGlu 1580
4712 TCCCTAGTAGATCCAACTCGGACCTCTGACAGAGAGGCTTCCCATCTCTCTGAA 4771
1581 SerSerThrGlyProSerTyrSerArgSer 1590
4772 AGTTCTACCGGACCTCTTACTCTCTGTTCA 4801

RESULT 3

AB021131
LOCUS Mus musculus RIM2 mRNA, 4800 bp mRNA linear ROD 20-DBC-2000
DEFINITION AB021131
ACCESSION AB021131
VERSION AB021131.1 GI:11611474
KEYWORDS Rim2.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (sices)
AUTHORS Ozaki, N., Shibasaki, T., Kashima, Y., Miki, T., Takahashi, K., Ueno, H., Sunaga, Y., Yano, H., Matsuura, Y., Iwanaga, T., Takai, Y. and Seino, S.

QY 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
DB 1232 AGGCACACATGAGAGAGGACAGTGTCTTTTGGCAACGCCTGAACCTAGAGAT 1291
QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
DB 1292 TCCAGGATTTCTGCTAAGGATGATAGACCATCAAGGCAAGATGCTATCTGAACGT 1351
QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
DB 1352 AGAGCTGCAATGGAAAACCAACGATCGTATTCAATGGAAAGAACTCGAGAGGCTCAGGGA 1411
QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProThrProArgArgSerPro 480
DB 1412 CAAGTCTTATCCCAAGAGGACCTCAATCATAGTCTCTCCACCCCTCGGGGGAGCCCT 1471
QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
DB 1472 ATACCGCTTGATAGACCAAGCATGAGCGCGCTGACTCCCTACGGAACAGCACCACTTA 1531
QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
DB 1532 GATCCAGCTCTGCTGTGAGGAAAAACGAAAGCGAGAAAAAATGGAACCATGTTAAGGAAT 1591
QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
DB 1592 GATTTTGGATTTAGACAGTCCAGTCCAGTCAAGTGGCGGCCCCCAAGGCGCTCATAAA 1651
QY 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAla 560
DB 1652 TCCAGAAGAGAGGTAATATGCGCCAGGTTCCTAGCAGCTCGAGGAGGAGCTGGCA 1711
QY 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
DB 1712 TCCACACCTCAGTATACAGCTGTGATGATGTGAGCTGGAAGCGCAGAGTGTGAGTGAG 1771
QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
DB 1772 AAAGGGGACAGTCAAAAGGAAAAAGAAAACTAGTGAGCAGGAGTTTGTGCGATCT 1831
QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
DB 1832 AACACAGGCTGTGAGAGCAAAAGAAAGAGATGACTATGTTGGCCACTTTTGGAGAG 1891
QY 621 AspLeuGluTyrSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 640
DB 1892 GATTTGGAATGCTCTGAGCTCAGATTAAGGACTCTGGGGTAGATACCTGTAGTAGCACA 1951
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
DB 1952 ACCCTTAAGGAGGATAGCCATAGTGATAGACCCCTGTGACCTGGCAGCCATCCAAA 2011
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
DB 2012 GATGGAGATCGCTAATTTGTCGTATTATTATTAATAGCGTTTAAAGATGGAGTGTA 2071
QY 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyLysMetThrGluSer 700
DB 2072 CCTCGAGACTCAGGAGCAATGCTGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA 2131
QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
DB 2132 GGTGCACTTTGTGCAATTTATACCAAGTAAAAAGGAAAGTTTGTAGCTATATCTGAGGA 2191
QY 721 HisLeuArgProGlyAspGluValLeuGluTyrAsnGlyValArgLeuLeuGlnGlyAlaThr 740
DB 2192 CATCTTAGACCGAGTGATGAGTCTTGGATGGATGGAGGCTATTGCAAGAGGCCACA 2251
QY 741 PheGluGluValTyrAsnIleLeuLeuGluSerLysProGluProGlnValGluLeuVal 760
DB 2252 TTTGAGGAAGTTTACACATTTATCTAGAAATCCAAACCTGAAACCAAGTTGAGCTTGT 2311

QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
DB 2312 GTTTCAGGCCAATTGGAGATATTCCTAGATACTGATAGACGATGACACAACTGGAA 2371
QY 781 SerSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
DB 2372 TCCAGTTCTAGCTCATTTGAATCTCAAAAAATGACCGCTCTTCTATATCCGTTACCTCA 2431
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
DB 2432 CCCATGAGTCTGCGATGCTGAGGAGTGTCCCGCAGTTCTTATCTGACAGCTTTCATA 2491
QY 821 LysLeuTyrPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
DB 2492 AAACCTATGGTTTGACAGGTTGGTCACAGTTGATAGTTACAAATTTGGGAGCAAGGAT 2551
QY 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
DB 2552 CTCCCTTCCAGGGAAGATGGGAGCCCAAGGAATCTTATGTTAAGATTTTACTCTCTTCCA 2611
QY 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysValThrLeuGluProLys 880
DB 2612 GATAGAAGTGATAAAAAATAGAGAAAGCAAAAAACAGTCAAGAAACCTTTTGGAAACCCAAA 2671
QY 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
DB 2672 TGGAAACAGACTTTTCATTTATTTCTCTGTCCACCGAGAGAAATTCCTGTGACGATGCTG 2731
QY 901 GluIleThrLeuTyrAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGlu 920
DB 2732 GAAATTAACCTTTGGGATCAAGCTAGAGTTTCGAGAAGAGAGAGCGAATTTCTTAGGAGAG 2791
QY 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspGluProHisTyrTyrLysLeuGln 940
DB 2792 ATTTTAATGGAATTTGGAACAGCTTGTCTAGATGATGAGCCGCACTGGTATTAAGCTGAG 2851
QY 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
DB 2852 ACCCATGATGCTCTCTCATTTGCCACTCCCTGCGCTTCCCCTATCTGCCCCGGAGGAG 2911
QY 961 LeuHisGlyGluSerProThrArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
DB 2912 CTCCATGAGAGAGAGCCCAACGCGAGGCTGCAAGGTCGAAGAGATAAAGTGACAGTGA 2971
QY 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
DB 2972 GTGTCTGACTACGACTCGAGGATGGCTGGAGTAGTGTGAGATTATCGACACAAATGCGC 3031
QY 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
DB 3032 CGCGATCTTCAAGGCTCAAGTTGTGCGTGCCAGAACCAAGTCATGTCATCAATCATTTGC 3091
QY 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
DB 3092 TCACCATCAGGCTCTCTCTCATCGATAGATGTTATAGGAGCAAGGTCATGTCGCT 3151
QY 1041 SerAlaProProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
DB 3152 AGTGGCCCTCTCTCTCAAGGAATGTGGAACAGGGGCAAGGAGGACGCTGCTACTGCGC 3211
QY 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
DB 3212 CATTAACAACATTTAGCCGAATGGATAGACACCGTGTGATGATGATGACCATCTCTTCA 3271
QY 1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
DB 3272 GATAGAGAC----- 3280
QY 1101 GluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAsp 1120
DB 3280 ----- 3280
QY 1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140

Db 3280 ----- 3280
Qy 1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
Db 3281 ----- AGGTCTTCACCTCGTACCGGCTCTGTCACACAAAGCCATCAAGTACTCCG 3331
Qy 1161 GlyThrGlyArgArgGlyValGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
Db 3332 GAAACAGACGAGGGGGCGACAGCTTCACAGCTTCACCAAAAGGGAACATTGGAGAGA 3391
Qy 1181 SerAlaMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrlsGlnVal 1200
Db 3392 AGTGCTATGGATATAGAGAGAGAAATCGCAATGAACTTAAACAATACAAACAGTA 3451
Qy 1201 AlaglySerAspProArgLeuGlnAspTyrlHisSerLysTyrlArgSerGlyTrpAsp 1220
Db 3452 GCCCGATCAGACCCCGACAGCTGAGCAAGATTACCATTCGAAAGTATCGCTCAGGATGGAT 3511
Qy 1221 ProHisArgGlyAlaAspThrValSerThrLysSerAspSerAspValSerAspVal 1240
Db 3512 CCACATAGGGGGGAGACTGTTTCCACTAAATCCTCGACAGTATGTAAGTATGTA 3571
Qy 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrlMetSerVal 1260
Db 3572 TCTCGGGTTTCAAGGACTAGTAGTCTTCTGTTTCAGACAGCAAGCTACATGTCCTC 3631
Qy 1261 GlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsn 1280
Db 3632 CAATCAGACGCGCCGAGAGAAACAGGAAATCAGTGTCTTTACATCCAAATGCAAAAC 3691
Qy 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerLysSerGlyAspMet 1300
Db 3692 AGACAGATGGGGCTGCGGGGAAGAACTGTACCAAAAGCACAGCATCAGTGGAGACATG 3751
Qy 1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320
Db 3752 TGCTCAGTGGAGAGAAATGACGGCAGCAGCTCCGACATGAGTGGGGCGCTCGGTACC 3811
Qy 1321 SerGlyLysLysArgArgSerSertleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
Db 3812 AGTGCAAGAGCGCGCATCTAGCATTTGGGGCAAAATGGTAGCTATTGTTGGTCTCTCA 3871
Qy 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArg 1360
Db 3872 CGGAAAGTCCGAGTCCCTCTCACTACGCAAAACCGAAGGAGGAGGTAAAGATACGG 3931
Qy 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
Db 3932 AGCACTGTTTCAGAGAGACCGGAGACCGGGCTAGCAGTGGAGATGAGGACCTGGATGACC 3991
Qy 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrlSerSerGluGlyAsn 1400
Db 3992 CGCCAGGCGCAGCGGGAATCCACAGATGCGACATGAACAGCTATAGCTCGAAGGAAT 4051
Qy 1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
Db 4052 CTGATCTTCCCTCGGGGTCGCGCTGACCTCTGACAGCCAGTTCAGTGAATTTCTCGATGGC 4111
Qy 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
Db 4112 CTGGCCCTCTCAGCTAGTGGAGCGGACAGCCCTGGCTACTCTCTGCAATGGGTGACAT 4171
Qy 1441 GlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleAArgAlaArgGly 1460
Db 4172 CAGTGGGAATGATGATAAAAGGACAGCTGGAGGTAGAAATCATTCGCGGCGCGCGC 4231
Qy 1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrlValLysValTyrlLeuLeu 1480
Db 4232 CTTGTGGTAAACACAGGTTTCAAGACACTGCGCAGCACCGCTATGTCAGGTGTATCTGTA 4291
Qy 1481 AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeuGluPro 1500

Db 4292 GACAACGGAGTCTGCATAGCCAAAAAGAAACCAAGGTGGCGAGAAAGACCCCTGGAGCCC 4351
Qy 1501 LeuTyrlGlnGlnLeuLeuSerPheGluGlnSerProGlnGlyArgValLeuGlnIleIle 1520
Db 4352 CTGTACCAGCAGCTCTTGTCTTCGAGGAGAGCCGCCAGGGGAGGGTGTACAGATCAT 4411
Qy 1521 ValTyrlGlyAspTyrlGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
Db 4412 GTCTGGGAGATATATGGTCGTATGATGATCAAAATCCCTTTATGGAGTGGCCACATACTC 4471
Qy 1541 LeuAspGluLeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPheProProSer 1560
Db 4472 TTAGATGAACCTGGACATATCCAAATGATGATGATGATGATGATGATGATGATGATGAT 4531
Qy 1561 SerLeuValAspProThrSerAlaProLeuThrArgAlaSerGlnSerSerLeuGlu 1580
Db 4532 TCCCTAGTAGATCAACCTTGGCACCTCTGCAAGAAAGAGCTTCCCAATCGTCTCTGAA 4591
Qy 1581 SerSerThrGlyProSerTyrlSerArgSer 1590
Db 4592 AGTTCTACCGACCTTCTTACTCTCGTTCA 4621

RESULT 4
AF199331
LOCUS AF199331 5550 bp mRNA linear ROD 04-JUL-2000
DEFINITION Rattus norvegicus RIM2-5B (Rim2-5B) mRNA, complete cds.
ACCESSION AF199331
VERSION AF199331.1 GI:8925875
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 5550)
AUTHORS Wang, Y., Sugita, S. and Sudhof, T.C.
TITLE The RIM/NIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins
JOURNAL J. Biol. Chem. 275 (26), 20033-20044 (2000)
MEDLINE 20347919
PUBMED 10748113

REFERENCE 2 (bases 1 to 5550)
AUTHORS Wang, Y. and Sudhof, T.C.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-1999) Center for Basic Neuroscience, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, TX 75235-9111, USA

FEATURES
Source
1..5550
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..5550
/gene="Rim2-5B"
52..4629
/gene="Rim2-5B"
/codon_start=1
/product="RIM2-5B"
/protein_id="AAF81653.1"
/db_xref="GI:8925875"
/translation="MSAPLGRGRPAATPAASQPPPPPEPDLHLTEERKIIOAVM
DRQKSEKSGSVLKLHQFMYTKQVKGWGESQOQBGDAFTCGIHKTKFPAD
CGHCNCSYCTKFCARCGVLSNKNVMVCMLCRQOILTKSGAFVYNGSNTQ
QPDQALRLRSEAPQKAKLHQFQPGDSSVPAVERGAHLTKRQDSIKNG
SMKHQIASDPSDRKSPSVSDONRERYDQSERREYSQVPSDSPTSPSPYADR
RSQEPQFPEPHLYRDSNRCHRSKEYIVDDDEVSDRDERGRREERYQART
SDPLARIPVQPIEYEMRIHAESVARSRRHSRHSVSLAHLSDLSRLNDRFS
RQDSRFRAMENQESYMERTEAQSSYPTQTNHSPPTPRPSPIPLDRPLER
ADSLRKQHLHLPSSAVTKRKEMTMLRNDLSLSDQSSVSRPPPHSKSGKMKR
QVSLSSSEELASTPEYTCDDVIESVGEKGMYSMLHASMSSEASPMSLHP
VTWQPSKQDRLIGRIILNKLKDGSPVPSGMLGLKVVGKMTESGRICAFITKVK
KGLADTVQHLRPGDELVNRLGATFEETVNIILSKRPQVSLVSRPGLDMP
RIPDTHAQLESSSSSSPESQMDRPSISVTPSPMSPGMLRDVQPLSQQLSIKLWFDKY

GHOLIIVTLGAKDLPSRGDRPNPYVKIYFLPDRSDKNKRRTTKVKKTLLEPKWNTF
 IYSPVRRERERLEITLDQARVRESEFEITLLELLEALLDDEPHWKLQTHD
 VSLPLPSPYMPRQLGCESTRRLQSKRI SDSSEVDYDCEDGCVGVYDRHGR
 DLGSLTSLVEQVMSNHCSPGPHRVDVIGTRSRWSVPPPPQRNVSGLGKTRAT
 GHYNTISRMDRHVMDDHYSSEDSHFLTLPRSHRQTSEHHRDGRDCEAADRQYH
 RSTSRQRLERTTTRSSRSRADTNLMSFPLMTGRSAPPSPALSRHPRTGSVQ
 TSPSSPTVTRRRQQLPQPPKGTLSRSMADIERNKQMLNKYKOVAGSDPRLQDY
 HSKYRSGMDPHRGADVTSTKSDSDVSAVSTSSASRFSSTSVMSVQSRPRGNR
 KLSVFTSKMQSRMGVSGKMAKSTSLSGDMCSLEKNGDSQSDTAVALGTSKRRS
 SIGAKMVALVGLSRKRSASQLSQTEGGGKKLRSTVQRSTETGLAVENNMTRQASR
 ESDTGNMSYSSSEGNLIFPGVRLASDSQSFDLGIGPAQLVGRQTLATPAMGDIQV
 MMDKQGLVEITRRAQLVVKPGSKTLPAPYVKVILLDNGVCIAKKTIVARKTLEPL
 YQGLLSFSPQKQLQIIVMGDYGRMDHKSFMGVAIILLDELELSNMVWGNPKLPP
 SSLVDPLAPLTRASOSSLSESTGPSYSRS"

BASE COUNT 1663 a 1281 c 1388 g 1218 t
 ORIGIN

Alignment Scores:

Pred. No.: 2,93e-226 Length: 5550
 Score: 7459.50 Matches: 1452
 Percent Similarity: 91.38% Conservative: 21
 Best Local Similarity: 90.07% Mismatches: 30
 Query Match: 90.04% Indels: 109
 DB: 10 Gaps: 4

US-09-617-099b-1 (1-1590) x AF199331 (1-5550)

QY 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 DB 52 ATGTCGGCTCCACTCGCGGCCCGCGGCGCGCGCTCCCAACCCCGCGCGCTCTCAGCCT 111
 QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIle 40
 DB 112 CCTCCGAGCCCGAGATGCCGACCTCAGCCACTCAGCAAGAGAGAGAGAGAGAAATCATC 171
 QY 41 LeuAlaValMetAspArgGlnLysLysGluGluGluLysGluGlnSerValLeuLysIle 60
 DB 172 CAGGCTGTCTATGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
 QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
 DB 228 ----- 228
 QY 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
 DB 229 -----AAG 231
 QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
 DB 232 CTGCATCAACAAATTTGAATGTATAGGAGCAAGTCAAGAAATGGGAGAGGAATCACAA 291
 QY 121 GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
 DB 292 CAGCAGCAAG 351
 QY 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
 DB 352 GCAGATGATGCGCGCATTAATGTTTATATGTCGCAACCAAGTCTGTGCTGTGTGGA 411
 QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
 DB 412 GGTGAGTGTATTACGCTCAACCAAGTTATGTGGGTGTGTATTTGTGCGCAACCA 471
 QY 181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
 DB 472 CAAGAAATCTCTACTAAGTCGGCGGTGTGTTTACAAATAGTGGGTCTAATACACCGCAG 531
 QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluAlaProGlnGluLysLys 220
 DB 532 CAGCCTGATCAAAAGGCTCTTCAGGAGCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAG 591
 QY 221 AlaLysLeuHisGluGlnProGlnPheGlnGlnValAlaProGlyAspLeuSerValProAla 240
 DB ----- 1686

DB 592 GCAAAACTGCATGAGCAGACGCGAGTTCCAGGAGACCCCGGTGACTCATCAGTACCTGCA 651
 QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
 DB 652 GTTGAGAGAGGTGCGACTCATGGCTCACAAGACAGGATCTTATTAAGAATGCTCAGGA 711
 QY 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 DB 712 ATGAAGCACAGATCCCGAGTGCATGCTTACAGACAAAAGAAAGTCCGTCAGTGTCC 771
 QY 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluLysGluLysArgLysValSerGlnTyrVal 300
 DB 772 AGGGATCAAAATCGAAGATACGACCAAGTGAAGAAAGAGAGAGATATTCACAGTATGTT 831
 QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320
 DB 832 CCTTCAGATAGCACAAATGCTAGATCTCCATCAGATTAATGCTCATAGACGATCTCAACGT 891
 QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgGly 340
 DB 892 GAACTCAATTTTATGAAGAACCTGATCAATTTAATTTATAGGATTTCTAACAGGAGAGC 951
 QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
 DB 952 CATAGACATTTCCAAAGAGTATTTGTAGACGACGAGATGTGAGACGACGAGATGAATAT 1011
 QY 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
 DB 1012 GAAAGACAAAGAGAGAGAGAGAGTACCAAGGACGCTACAGAAAGTCAATCAAAATTTGGCC 1071
 QY 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
 DB 1072 CGGTATCCGTTAAGCCACACCTTATGAAGACGAAATCGGATCCACCTCAAGTGTCC 1131
 QY 401 ArgAlaArgHisGluArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
 DB 1132 CGGCGACGACACAGAGAGAGGACAGTGAATGTTCTTGGCAATGCTGAATAGAGAT 1191
 QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
 DB 1192 TCCAGGATTTCTCTGCTAGGATGATAGACCATCAAGGCAAGATCTGATCTCAACGT 1251
 QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
 DB 1252 AGAGCTGCAATGGAACCCAGCGATGCTATTCATGGAAGAACTCGAAGACTCAGGA 1311
 QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
 DB 1312 CAAAGTTCTTATCCCAAGGACCAAAATCATAGTCTCTTACCTCCCTCGCAGAGACCT 1371
 QY 481 IleProLeuAspArgProAspMetArgAlaAspSerLeuArgLysGlnHisLeu 500
 DB 1372 ATACCGCTCGATAGACGAGCTGAGCGGTGCGGACTCCCTACGGAACCAACACCACTTA 1431
 QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgan 520
 DB 1432 GATCCAGCTCTGCTGTAAAGGAAAACGAGCGAGAAAAAATGGAACAAATGTTGGAAT 1491
 QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProProHisLys 540
 DB 1492 GATTTCTTGGATTACAGCAGTCTGAGTCAGTGAGCGCCGCCCCACCAAGGCTCATAA 1551
 QY 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560
 DB 1552 TCCAGAAAGAGAGGTAATAATGCGCAGGTTTCACTAGCAGCTCAGAGAGAGGTTGCA 1611
 QY 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerSerGluSerValSerGlu 580
 DB 1612 TCCAGCTCTGATATACAGCTGTGATGACGTGGAGATTGAAGCGAGAGCGTAGTGAG 1671
 QY 581 LysGlyAspSerGlnLysLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
 DB 1672 AAGAGACATGAG----- 1686

QY 601 AenThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyHisSerLeuGluGlu 620
Db 1687 -----TACAGCTGGTGGAG 1701
QY 621 AspLeuGluTrpSerGluProGlnLysLysAspSerGlyValAspThrCysSerSerThr 640
Db 1702 CATGGCTCTGGCAT----- 1716
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1717 -----AGCAGTGAGCATCCCAAGTCTTTGGACCCCTGTGACCTGGCAGCCATCCAAA 1770
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 1771 GATGAGATCGCTAATTTGGTGGTATTTTAAATAAGCGTTTAAAGATGGAGTGA 1830
QY 681 ProArgAspSerGlyAlaMetLeuGlyLysValValGlyGlyLysMetThrGluSer 700
Db 1831 CTTCCAGATTCAGGAGCAATGCTGGCTTAAAGGTTTAAAGGTTTAAAGTGGAGTGA 1890
QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 1891 GGTGACCTTTGGCTATTTATACCAAGTGAAGGAAAGGAGTTTACGTGATCTAGGA 1950
QY 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 1951 CATCTTAGACAGGTGATGAAGTCTTGGATGGATGGAGCTATTGCCAGGACCCACA 2010
QY 741 PheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVal 760
Db 2011 TTTGAGGAAGTTTACAACATTTATCTAGAAATCCAGGCTTGAACCAAGTTGAGCTTGT 2070
QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2071 GTTTCAGGCCGATTGGAGATAGCTTGAATACCTGATAGACATCATGCAACATCGGAA 2130
QY 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2131 TCCAGTTCTAGCTATTGATCTCAAAAATGGATCGTCCGTCTATATCGTTACCTCT 2190
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2191 CCCATGAGTCTGGCATGCTGAGGAGATGTCACAGTCTTATCTGGACAGCTTTCAATA 2250
QY 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2251 AAACATATGGTTTGACAGGTTGGTCCACAGTTAATAGTTACAAATTTTGGGAGCAAGGAT 2310
QY 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2311 CTCCCTCCAGGAGATGGGAGCCAGGAATCTTATGTTAAATTTACTTCTTCCA 2370
QY 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2371 GACAGAAGTGATAAAAAACAGAGAAGAACAAAAACAGTCAAGAAAACTTTGGAAACCCAAA 2430
QY 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2431 TGGAAACAGACTTTTCATTTATTTCTCTCTTCCCGAAGAGAAATTCGCGGAAACGAAATGCTC 2490
QY 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGlu 920
Db 2491 GAATCACCTTTGGGATCAAGCTCGAGTTCCAGAGGAAGAAAGTGAATTTCTAGGAGAG 2550
QY 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGln 940
Db 2551 ATTTTAAATGAATGGAAACAGCTCTGTATGATGATGATGATGATGATGATGATGATGAT 2610
QY 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
Db 2611 ACCCATGATGCTCTCTCAATTCCTCCTCCCTCCTCCATATATATGCTCCAGGAGACAG 2670

QY 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db 2671 CTGCATGAGAGAGAGCCACACGAGTTTCAAGGTCGAGAGGANTTAAGTACAGCGAA 2730
QY 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
Db 2731 GTATCCGACTATGACTGGAGAGCGCGTGGAGTGGTGTGAGATTACCGACACCATGGC 2790
QY 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
Db 2791 CGTGATCTTCAAGCTCCACATTTACGTGCCAGAACCAAGTCATCTCATCAACCATTTGC 2850
QY 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
Db 2851 TCACCATCAGGCTCTCTCATCGAGTAGATTTATAGAGAGACTAGGTCTATGGTGGCT 2910
QY 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 2911 AGTGTCCCTCTCTCTCAAGGAAATGTGAAACAGGGCTTCGAGGACACAGTGTACTGGC 2970
QY 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 2971 CATTACAATCAATTTAGCCGATGGACAGACACCGTGTCTATGGACGACCATTTACTCTTCA 3030
QY 1081 AspArgAsp----- 1083
Db 3031 GAGAGAGACATCTTTCTCCT 3090
QY 1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
Db 3091 CACCACAGGAGTGAAGGATTTGTGAAGCAGACAGATAGACAGCCATATCAGATCCAGA 3150
QY 1099 SerThrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArg 1118
Db 3151 TCAACAGAACCAACGGCTCTCTCTAGAGCGGACCAACCCGCTCCAGATCTCTCTGAAGCT 3210
QY 1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
Db 3211 GCTGATCAAACTCTATGAGGTGATGCTTATTAATGATGGAAGATCTGCCCCCTCT 3270
QY 1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
Db 3271 TCACCTGCTTATCGAGGTCTCACCTCGCATGGCTCTGTCCAGACCAAGTCCGTTCAAGT 3330
QY 1159 ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeu 1178
Db 3331 ACTCCCTTAACAGAGCGGCGGCGGAGCTTCTCTCAGCTCCCAACCAAGGAAACACTG 3390
QY 1179 GluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLys 1198
Db 3391 GAAGAAGTGTCTATGGATATAGAGAGAGAAATCGCAAAATGAAACTTAAACAAATACAAA 3450
QY 1199 GlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysValTyrArgSerGly 1218
Db 3451 CAGGTAGCCGATCAGATCCAGACTGGAACAGATTAACATTCGAGATTCGTTCTCAGGA 3510
QY 1219 TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSer 1238
Db 3511 TGGGATCCACATAGAGCGGAGATCTGTTTCTACTAAATCTCTCGGACAGTGAATGAAT 3570
QY 1239 AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMet 1258
Db 3571 GATGTATCTCGGTTTCAAGGACTAGTAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3630
QY 1259 SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet 1278
Db 3631 TCCGTCCATCAGAACGCCGAGGAGGAAACAGGAAATCAGTGTCTTTAGTCTCAAAATG 3690
QY 1279 GlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGly 1298
Db 3691 CAAAGCAGACAGATGGGTGTGTGGGGAGAGCATGCCAAAAGCACCAGCATCAGTGA 3750
QY 1299 AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu 1318

Db 3751 GACATGCTGCTACTGAGAGATGATGCGAGCCAGTCGGACACTCGAGTGGGCGCCCTG 3810
 Qy 1319 GlyThrSerGlyLysArgSerSerIleGlyAlaLysMetValAlaLysValGly 1338
 Db 3811 GGCACAGTGGCAAGAGCGCGATAGCATTTGGGCGCAAAATGGTAGCTATTGTGTGT 3870
 Qy 1339 LeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLys 1358
 Db 3871 CTCTCAGGAAGAGTGGAGTGGCTCCCACTCAGCCAAACCGAGAGAGGTAAG 3930
 Qy 1359 LeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAenTrp 1378
 Db 3931 CTAGCGAGCACTGTCAGAGAGACGAGAGCGGCTAGCACTGGAGATGAGACTG 3990
 Qy 1379 MetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerGln 1398
 Db 3991 ATGACAAGACAGCGGCGGAAATCCAGGATGGCAGCATGAACAGTTACAGCTCAGAA 4050
 Qy 1399 GlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeu 1418
 Db 4051 GGAATCTGATCTTCTGGGGTCCGCTGCGCTCTGACAGCCAGTTCAGCGATTTCCT 4110
 Qy 1419 AspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly 1438
 Db 4111 GATGTCCTGGCCCTGCTCAGCTAGTGGAGCGCCAGACCTGGCGACTCTCGCAATGGT 4170
 Qy 1439 AspIleGlnValGlyMetMetAspLysGlyGlnLeuValGluIleLeuArgAla 1458
 Db 4171 GACATTCAGTGGGGATGATGGAAGAAGAGGACAGCTGGAGTAGAGATCATCGGGCT 4230
 Qy 1459 ArgGlyLeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyr 1478
 Db 4231 CGAGGCTTGTAGTAAACAGGTTCAAGACACTCCAGCACCATTATGCAAGGTGAC 4290
 Qy 1479 LeuLeuAspAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeu 1498
 Db 4291 CTGTTAGACAACGGAGTCTGCATAGCAAAAGAAACCAAGCTGGCGAGAAAGACGCTG 4350
 Qy 1499 GluProLeuTyrGlnGlnLeuSerPheGluSerProGlnGlyArgValLeuGln 1518
 Db 4351 GAGCCCTGTATCAGCAGCTGTTATCTTCGAAGAGAGCCCGCAGGAGGAGTGTACAG 4410
 Qy 1519 IleLeuValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGln 1538
 Db 4411 ATCAITGCTGGGAGATATGCGCGCATGGATCAACAATCTCTTATGGAGTGGCCAG 4470
 Qy 1539 IleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPhePro 1558
 Db 4471 ATACTCTTAGAACTGGAACCTATCCAACTGTAATGGATGTTCAAACTCTTCCCT 4530
 Qy 1559 ProSerSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSer 1578
 Db 4531 CCTTCTCTCCTAGTAGATCAACCTTGGCAGCTCTGACAGAGAGAGCTTCCCAATCGCT 4590
 Qy 1579 LeuGluSerSerThrGlyProSerTyrSerArgSer 1590
 Db 4591 CTGGAAGTCTACTGGAACCTTCTTACTCTCGTCA 4626

RESULT 5
 AF199330 5592 bp mRNA linear ROD 04-JUL-2000
 LOCUS Rattus norvegicus RIM2-5A (Rim2-5A) mRNA, complete cds.
 DEFINITION AF199330
 ACCESSION AF199330.1 GI:8925873
 VERSION
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE 1 (bases 1 to 5592)
 AUTHORS Wang, Y., Sugita, S. and Sudhof, T.C.

TITLE The RIM/NIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins J. Biol. Chem. 275 (26), 20033-20044 (2000)
 JOURNAL 20347919
 MEDLINE 10748113
 PUBMED
 REFERENCE 2 (bases 1 to 5592)
 AUTHORS Wang, Y. and Sudhof, T.C.
 TITLE Direct Submission
 JOURNAL Submitted (27-OCT-1999) Center for Basic Neuroscience, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, TX 75235-9111, USA
 FEATURES
 Location/Qualifiers
 1..5592
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..5592
 /gene="Rim2-5A"
 52..4671
 /gene="Rim2-5A"
 /codon_start=1
 /product="Rim2-5A"
 /protein_id="AAF81652.1"
 /db_xref="GI:8925874"
 /translation="MSAPLGRGRPAPTPAASQPPPOPEMDLSLTLTEERKIIQAVM DROKKEEKQSVLKLHQQFEMVKQVKMGESQOQEQKQDAPTCGICHTKFPAD QDHCNVCQKFCACGCGYSLSNKVMWNCNLCKQKEILTKSGANFYSGSNTPQ QPDVALAGLSERAPQKAKLHEQTQFGPPGDSVPAVERGARGLTQDSIKNG SGMKQIASDMPSDRKRSPSDQNRRIYDQSEERYSQYVPSDSTMPSPSYADR RSQEPQYEPEDHLNTRNRGRHSKEYIVDDVEDSDRDEYERQREBEYQARYR SDPLNARYPVKPYEQMRIAHSVRARRHRRSDVLANAELEDSPIILMDRPS RQSRYSERRAAMENORSYMERTEAQCQSSVPQRTTNHSPTPERSPIPLDRPELR ADSLKHLDLDPSSAVALKEKMETMLRNLSLSDQSESVRPPPHKSKGGMKR QVLSSESEELASTPYTSCDDVEISESGVGMDEYSWLEHASWHSSESPMSLHP VTQPSKDDDELIGILLNKLKQDSVPRDSGAMGLGVGCKMTESGRLCAFTTKVK RGSADITVGLHPGDEVLENGRLIQCATFEVNIILLESKPEQVLEIVSRPIGDMF RPDSTHAQLBSSESPESQKMDPSISVTSFMSQMLRDVDPVPSGOLSVKLMFQV GHOLIVTILGAKDLPSREDGPRNPYKVIYLPDRSDKNKRTKTKVTLKPKWNTF IYSPVRRERFERMLEITLMDQARVRESEFGLLELLELLELLELLELLELLELLEL VSSLLPSPSPMPFAPQHGSPTRLQKRSKISDSSEVSDYDCDGVGVSDYRHDGR DLQSLTSLSPVQMSNNHCSFSPGPHRVIVTGRTSFSPVPPQNTYQGLRGTRAT GHYNTISMDRHRVMDHYSSERDSHFLTLPSSRHROTSEHHRDRCCEAADROPYH RSRSTEORPILLERTTTRSRSERADTNLMRSPSLMTGRSAPPSPALSRSHPTGVSQ TSPSTPYTGERGRLPOLPKGTLEMITEDMSTRKNSGAMDIERNQMKLANKY KQVAGSDPRLSQDHSKTRSGMDPHRGADTVTSKSDSDSDSVSAVSSTSSASRFSST SYMSVQSRSPRGRNKISVFTSKMOSRQMGVSGKMAKTSISGDMCSLEKNDQSQSDT AVEENWMTROASRESTDGSNNYSSEGNLIPFGVRLASDSQFDFDLGLGPAQVGR OTLATPAMGDIQVGMMDKKGLEVEIIRARGLVKPKSKTLPAPYVYVLLDNGVCLIA KKTKVARKTLEPLYQQLLSPEESQGVLOIIVMGDYGRMDHKSFMGVAQILLDELE LSNVIGWFKLFPSSSLVDPTLAPLTRASQSSLESSTGSPSYRS"

BASE COUNT 1681 a 1290 c 1398 g 1223 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,35e-225 Length: 5592
 Score: 7438.50 Matches: 1451
 Percent Similarity: 90.53% Conservative: 21
 Best Local Similarity: 89.24% Mismatches: 31
 Query Match: 89.78% Indels: 123
 DB: 10 Gaps: 5

US-09-617-099B-1 (1-1590) x AF199330 (1-5592)

Qy 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 Db 52 ATGTCGGCTCCACTCGGGCCCCGGCGCGCTCCACCCCGCGGCTCTCAGCT 111
 Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIlele 40
 Db 112 CCTCCGCGAGCCGAGATCGGACCTCAGCACCTCAGCGAGAGAGAGAAATCATC 171
 Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluGluLysGluGlnSerValLeuLysIle 60

781 SerSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
2131 TCCAGTCTTAGCTCAATTCGAAATCCAAAAATGGATCGTCCGCTATATATCGGTACCTCT 2190
801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
2191 CCCATGAGTCTCGGATGCTGAGGGATGTCACACAGTCTCTATCTGGACAGCTTTCAATA 2250
821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
2251 AAACATATGTTTGCACAGGTGTGTACACAGTTAATAGTTACAAATTTTGGAGCCAAAGGAT 2310
841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
2311 CTCCTCTCCAGGCAAGATGGAGGCCCAAGGAATCTTATGTATAAATTTACTTCTCTTCCA 2370
861 AspArgSerAspLysAsnLysArgThrLysThrValLysLysThrLeuGluProLys 880
2371 GACAGAAGTGATATAAACAGAGAGAACAAAAACAGTCAAGAAACCTTTGGAAACCCAAA 2430
881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
2431 TGGAAACCACTTTCATTTATTTCTCTGTTCCCGAGAGAAATTCGGGAAACGAATGCTC 2490
901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGlu 920
2491 GAAATCACCTCTTGGGATCAAGCTCGAGTTCGAGAGAGAAAGTGAATTCCTAGGAGAG 2550
921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGln 940
2551 ATTTTAATGAATGGAAACAGCTCTGTAGATGATGAACCAACACACCTGGTACAAATTCAG 2610
941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgGln 960
2611 ACCCATGATGTCTCTCATTTGCCACTCCCTCACCTTCTCCATATATGCCAGGAGACAG 2670
961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
2671 CTGCATGGAGAGAGCCCAACAGCAGGTTGCCAAGGTCGAAGAGTCAAGTAAGTGACACGCAA 2730
981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
2731 GTATCCGACTATGACTGCCAGAGACGCGCTGGAGTCTGTCTCAGATTACCGACACGATGCG 2790
1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
2791 CGTGATCTTCAAAGCTCCACATTTATCAGTCCAGAACAAAGTCAATGTCATCAACCACTTGC 2850
1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
2851 TCACCATCAGGCTCTCTCATCGAGTAGATGTTATAGGAAGGACTAGGTCTATGTCGCT 2910
1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
2911 AGTGTCCTCTCTCTCAAAGGAATGGGAACAGGGGCTTCGAGGGACACAGTGTCTACTGCGC 2970
1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
2971 CATTACAAATCAATTAAGCCGGATGGACACACCCGTTGATGGACCAACCACTACTCTTCA 3030
1081 AspArgAsp----- 1083
3031 GAGAGAGACAGTCATTTTCTCACTTACTCTCGCTCTCCGACACAGGACAGCAGTGAGCAT 3090
1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
3091 CACCACAGGGATGGAAAGGATTTGTGAAGCAGCAGATAGACAGCATATCACAGATCCAGA 3150
1099 SerThrGluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArg 1118
3151 TCAAACAGAACACGGGCTCTCTCTAGAGCGGACCAACCCGCTCCAGATCCTCTGAACGT 3210

1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
3211 GCTGATACAAACCTCATGAGTTCGATGCTTTCTTAATGACTGGAAAGATTCGCCCTCTCT 3270
1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
3271 TCACCTGCTTATCGAGGTCTCACCTCGCACTGGCTCTGTCCAGACCAAGTCCGTCAGAT 3330
1159 ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeu 1178
3331 ACTCCCGTAAACAGGAGCAAGCGCGCGAGCTTCTCTCAGCTCCCAACAAAGGAAACACTG 3390
1179 GluArg-----SerAlaMetAsp 1184
3391 GAAAGATGATCACAGAAACATGAGCTCCCAAGAAAGAGAAACTCAGGTGCTATGGAT 3450
1185 IleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGlySerAsp 1204
3451 ATAGAGGAGAGAAATCGCCAAATGAACTTAACTTAACAAATCAACAGGTAGCCGATCAGAT 3510
1205 ProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTyrAspProHisArgGly 1224
3511 CCCAGACTGGAAACAGATTCACATTCGAAGTATGCTCAGGATGCGATCCACATAGAGGG 3570
1225 AlaAspThrValSerThrLysSerSerSerSerSerSerSerSerSerSerSerSerSer 1244
3571 CGAGATACCTGTTCTACTATAATTCCTCGGACAGTGAATGTAAGTATGATATCTGCGGTTCA 3630
1245 ArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerValGlnSerGluArg 1264
3631 AGGACTAGTAGTCTCTCGTTTCAGCAGACACAGTACATGTCCTCATTCAGAACCG 3690
1265 ProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsnArgGlnMetGly 1284
3691 CCGAGAGAAACAGGAAATCAGTGTCTTACGTCCAAATGCAAAAGCAGACAGATGGGT 3750
1285 ValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMetCysSerSerLeuGlu 1304
3751 GTGTCGGGAAAGACAGTGGCCAAAAGCACCAGCATCAGTGGAGACATGTGCTCACTGGAG 3810
1305 LysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLys 1324
3811 AAGAATGATGGCAGCAGTGGACACTGAGTGGGCGCCCTCGGCGCACAGTGGCAAGAG 3870
1325 ArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLysSerArgLysSerArg 1344
3871 CGGCGATCTAGCATTTGGGCGCCAAATGCTAGCTATTGTTGTTCTCTCACGGAAGATCGC 3930
1345 SerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLysLeuArgSerThrValGln 1364
3931 AGTGCTCCCAACTCAGCCAAACCGAAGGAGGAGGTAAAGACTACGAGCACTGTCCAG 3990
1365 ArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSer 1384
3991 AGAAGACAGGAGACCGGCTAGCAGTGGAGATGAGAACTGGATCAACAGACAGCCACAGC 4050
1385 ArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyLysLeuIlePhePro 1404
4051 CGGGAATCCACGATGGCAGCATGAACAGTTACACTCAGAGGAAATCTGATCTTTCTCT 4110
1405 GlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGlyProAla 1424
4111 GGGGTCCGCTCGGCTCTCAGCAGCCAGTTCCAGCGATTTCTTGTATGTTCTGCGGCTGT 4170
1425 GlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMet 1444
4171 CAGCTAGTGGGAGCGCCAGACCTCTGCGACTCTCTCAATGGGTGACATTCAGGTGGGAGATG 4230
1445 MetAspLysLysGlyGlnLeuGluValGluIleLeuArgAlaArgGlyLeuValLys 1464
4231 ATGGATAGAGAGGAGCAGCTGGAGTAGAGATCATCCGGGCTCGAGGCTTGTAGTAAAA 4290
1465 ProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeuAspAsnGlyVal 1484

```

Db      4291  |||||CCAGGTTCCAGACACTGCCAGCACCATTATGCTCAAGGTGTACCTGTTAGCAACGCGAGTC 4350
Qy      1485  CyrllealalyelyleThrlleValalaArglyeThrlleuGluProLeuTyrlGlnGln 1504
Db      4351  |||||TGATAGCCAAAGAAAGAAACCAAGGTGGCGAGAAAGCGCTGGAGCCCTGTATCAGCAG 4410
Qy      1505  LeuLeuSerPheGluGluSerProGlnGlyArgValleuGlnllelleValtTpGlyAsp 1524
Db      4411  |||||CTGTTATCTTTCGAAGAGAGACCCCGAGGAGGTGTACAGATCATGTGTCGGGAGAT 4470
Qy      1525  TyrlGlyArgMetAspHisLysSerPheMetGlyValalaGlnlleleuLeuAspGluLeu 1544
Db      4471  |||||TATGGCGGATGATGATCAAAATCTTTATGGAGTGGCCAGATACTCTTAGTAGAAGT 4530
Qy      1545  GluLeuSerSerMetValilleGlyTrrPheLysLeuPheProProSerSerLeuValasp 1564
Db      4531  |||||GAATATATCAACATGGTAATGGATGTTCAAACTCTTCCTCTCTCTCTCTCTCTCTCTCT 4590
Qy      1565  ProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGluSerSerThrGly 1584
Db      4591  |||||CCAACTTGGCACTCTGACAGAGAGAGCTTCCCAATCGTCTCTGGAAGTCTTACTGGA 4650
Qy      1585  ProSerTyrlSerArgSer 1590
Db      4651  |||||CTTCTTACTCTCGTTCA 4668

RESULT 6
AF199322 5640 bp mRNA linear ROD 04-JUL-2000
LOCUS The Rattus norvegicus RIM2 (Rim2) mRNA, complete cds.
DEFINITION AF199322
ACCESSION AF199322.1 GI:8925857
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 5640)
AUTHORS Wang,Y., Sugita,S. and Sudhof,T.C.
TITLE The RIM/MIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins
JOURNAL J. Biol. Chem. 275 (26), 20033-20044 (2000)
MEDLINE 20347919
PUBMED 10748113
REFERENCE 2 (bases 1 to 5640)
AUTHORS Wang,Y. and Sudhof,T.C.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-1999) Center for Basic Neuroscience, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, TX 75235-9111, USA
FEATURES
source
    1..5640
        /organism="Rattus norvegicus"
        /mol_type="mRNA"
        /db_xref="taxon:10116"
gene
    1..5640
        /gene="Rim2"
        /codon_start=1
        /product="RIM2"
        /protein_id="AA081644.1"
        /db_xref="GI:8925858"
CDS
    52..4719
        /gene="Rim2"
        /codon_start=1
        /product="Rim2"
        /protein_id="AA081644.1"
        /db_xref="GI:8925858"

```

```

ADSLRKHQLDPSSAVRKTREKMETMLRNDLSLSDQSSSVRPPPPRPHKSKKGGKVR
QVLSSESEELASTPEYTSDDVIESSESGKGMDEYSLWLEHASWHSSEASMSLHP
WTPQSKDGLRLIGILLNKLKDGSPRDSGAMLGLKVGSGMTVSGRLCAPITTVK
KSLATDVCGLRPGDEVLWNGRLAGATFEVYNTLILSKPEQVLYVSRIGDMP
RIPDSTHAQLESSESSSFQKMDRPSISVTSFMSPGMLDVPQFLSGQLSSQSLSRRT
TFPVPVQIKLWPKVGHQLIVTILGAKDLPSRGRPNPYVKIYIFLDRSDKNKR
TKTVKTLSPKWNQTFIYSPHRRERFRLMLITLMDQARVEESESFLGELILELET
ALLDGVSDYRHDGRDLOSSTLSVPQVMSNHCSPPSRHVRDVIQTRWSVSPV
CEDGVSDYRHDGRDLOSSTLSVPQVMSNHCSPPSRHVRDVIQTRWSVSPV
PQORNVCEALDQPHRSTRSTORPILLETTRSRSERADTLMKMSMLMTGRSAP
HDDRCSEALDQPHRSTRSTORPILLETTRSRSERADTLMKMSMLMTGRSAP
HPALSRHPTGTSVTSFSTVTRGRQPLPQPKTLMETIMDSMTKRNNG
ANDIDERNQMKLNTKYVAGSDPRLEQDYHSKTSMDPHRGADTVTSKSDSDVS
VSAVSTSGASRFSSTYSVQVOSRPNRKRISVFTSKMQSRMGVSGKSKMAKSTIS
GMCLEKNDGSDTAVCALATSGKRRSSICAKMVAIVGLSRSRASQLSQTSRQ
GKLRSTVORSTETGLAVENRMTQASRETDGSMNSYSEGNLIPQVRLASRQ
FSDFLGQLPAQLGQTLATPAMGDIQVMDKQGLVEIIRAGLWVVKPSKTLTP
AFYFVGLDNGVCIATKKTIVARKTLEPLYQQLSFSPESPOQKVIQIIVMGDGRMD
HKSFMVQAQILLDELSNNVIGFELFPSSLVDPDLAPLTRASQSSLESSTGPSY
SRs"
BASE COUNT 1695 a 1301 c 1409 g 1235 t
ORIGIN
Alignment Scores:
Pred. No.: 5e-225 Length: 5640
Score: 7420.50 Matches: 1451
Percent Similarity: 89.65% Conservative: 21
Best Local Similarity: 88.37% Mismatches: 31
Query Match: 89.57% Indels: 139
DB: 10 Gaps: 6
US-09-617-099B-1 (1-1590) x AF199322 (1-5640)
Qy 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
Db 52 ATGTCGGCTCCACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCCAC 111
Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysLeile 40
Db 112 CTTCCGAGCCCGAGATGCGCGACCTCAGCACCTCAGGAGAGAGAGAGAGAGAGAGATC 171
Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluGluGluGluGluGluGluGluGlu 60
Db 172 CAGGCTGTCATGATCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 228 ----- 228
Qy 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGlnLysGluProGlnThrLys 100
Db 229 -----AAG 231
Qy 101 LeuHisGlnGlnPheGluMetTyrlLysGluGlnValLysLysMetGlyGluGluSerGln 120
Db 232 CTGCATCAACAATTTGAATGTATAGGAGCAAGTCAAGAAATGGAGAGAGATCAACA 291
Qy 121 GlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
Db 292 CAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 351
Qy 141 AlaAspGlyCysGlyHisLysCysSerTyrlCysGlnThrLysPheCysAlaAlaGly 160
Db 352 GCAGATGGATGGCGGCATACCTTTCATATGTCACCAAGTCTGTGTCTGTGTGTGGA 411
Qy 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysAsnGlyGln 180
Db 412 GGTGAGTGTTCATACCTCAACACAGATTAATGTGGTGTGTGTGTGTGTGTGTGTGGA 471
Qy 181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrlAsnSerGlySerAsnThrLeuGln 200
Db 472 CAAGAAATCTCTACTAAGTCGGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 531

```

QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGlnLysLys 220
DB 532 CAGCCTGATCAAAAGGCTCTTCAGGGCTTCGAGTGAAGAGCCCTCAGGAGAGAG 591
QY 221 AlaLysLeuHisGlnGlnProGlnPheGlnGlnAlaProGlyAspLeuSerValProAla 240
DB 592 GCANAACTGCATGACGACGAGTTCAGGACCCCGGTGACTCATCACTACCTGCA 651
QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
DB 652 GTTGAGAGAGTCCAGCTCATGGCTCAACAGACAGGATTCATTAAAGAAATGGCTCAGGA 711
QY 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
DB 712 ATGAGACACCATGCCAGTGCATGCCCTTCAGACAGAAAGAGTCCGTCAGTGCC 771
QY 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
DB 772 AGGGATCAAAATCGAAGATACGACCAAGTGAAGAAAGAGAGGAAATATTCACAGTATGTT 831
QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
DB 832 CCTTCAGATAGCACATGCTAGATCTCCATCAGATTATGCTCATAGACGATCTCAAGCT 891
QY 321 GlnProGlnPheTyrGluGlnProGlyHisLeuAsnTyrArgAspSerAsnArgGly 340
DB 892 GAACCTCAATTTATGAAGAACTGATCATTTAAATATAGGGATTCCTAACAGAGAGGC 951
QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
DB 952 CATAGACATTCNAAAGATATATTGTAGACGCGAAGATGTGGAGACGACAGATGAATAT 1011
QY 361 GluArgGlnArgArgGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
DB 1012 GAAAGACAAAGAGAGAGAGAGTACCAGGACGCTACAGAAAGTATCAAAATTTGGCC 1071
QY 381 ArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSer 400
DB 1072 CGGTATCCGGTAAAGCCACACCTATGAAGAGCAAAATCGGATCCACGCTGAAGTGCC 1131
QY 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
DB 1132 CGGGACGACACAGAGAGAGGACAGTATGTTCTTTGGCAATGCTGAATAGAGAT 1191
QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
DB 1192 TCCAGGATTTCTGCTAAGATGGATAGACCATCAAGCAAAAGATCTGTATCTGAAGCT 1251
QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
DB 1252 AGAGCTGCATGGAAACACAGCGATCGTATTCAATGGAAGAACTCGAAGCTCAGGGA 1311
QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
DB 1312 CAAAGTTCTTATCCACAAAGGACCAAAATCATAGTCTCTCCCTCCGAGGAGCCCT 1371
QY 481 IleProLeuAspArgProAspMetArgAlaAspSerLeuArgLysGlnHisHisLeu 500
DB 1372 ATACCGCTCGATAGACGAGCTGAGGCGTGCAGCTCCCTACGGAACCAACACCACTTA 1431
QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
DB 1432 GATCCAGCTCTCTCTGTAAAGGAAACGAAAGCGAGAAATAATGGAACAATGTTGAGGAAT 1491
QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
DB 1492 GATCTTTGAGTTACAGACAGTCTGAGTCACTGAGGCGCCGCCCAACAGGCTCATAAA 1551
QY 541 SerLysLysGlyLysMetArgGlnValSerLeuSerSerSerSerGluGluLeuAla 560
DB 1552 TCCAGAAAGAGGATTAATGCACCAGGTTTCACTGAGCAGCTCAGAGGAGGAGTTGGCA 1611
QY 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580

DB 1612 TCCAGCCTGAGTATACAGCTGTGATGACCTGAGATTAAGACGAGAGCTAGGTGAG 1671
QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlnValLeuSerAspSer 600
DB 1672 AAAGGAGACATGGAG----- 1686
QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
DB 1687 -----TACAGCTGGTGGAG 1701
QY 621 AspLeuGluTyrSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
DB 1702 CATGGCTCTGGCAT----- 1716
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTyrGlnProSerLys 660
DB 1717 -----AGCAGTGAGGCATCCCAATGCTTTGCACCCCTGTGACCTGCGACCTGCAGCACCATCCAAA 1770
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
DB 1771 GATGAGATCCCTAAATTTGGTCTGATTTTAAATAAGCGTTTAAAGATGGGAGTGA 1830
QY 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
DB 1831 CCTCGAGATTGAGGCGAATGCTGGCTTAAAGTTTGAAGGAAAGATGACTGAATCA 1890
QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
DB 1891 GGTGCGATTTGTGATTTATTATTAAGTAAAAAGAAAGTGTAGCTGATCTGTAGCA 1950
QY 721 HisLeuArgProGlyAspGluValLeuGluTyrAsnGlyArgLeuLeuGlnGlyAlaThr 740
DB 1951 CATCTAGACAGGATGATGAAGTCTGGAATGGAATGGAAGCTATTCCAGGAGCCACA 2010
QY 741 PheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVal 760
DB 2011 TTTGAGGAGTTTACAACTATTCTAGAAATCCAGCTGAAACCAAGTTGAGTTGTT 2070
QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
DB 2071 GTTTCAAGCCGATTTGGAGATATGCTAGATATCTGATAGTACCTCATGCACTCATGCA 2130
QY 781 SerSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
DB 2131 TCCAGTTCTAGCTCATTTGAATCTCAAAAAATGGAATCGTCCGCTCATATATCGGTACT 2190
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSer--- 819
DB 2191 CCATGATGCTCTGGCATGCTGAGGATGTCCACAGTTCTTATCTGGACAGCTTTCAGC 2250
QY 820 -----IleLysLeuTrpPhe 824
DB 2251 CAAAGCCTTAGTAGAAGAACCAAGCCTTTTGTCTCTAGGTTTCAGATAAAACTATGGTTT 2310
QY 825 AspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArg 844
DB 2311 GACAGGTTGGTGCACCGCTAATAGTTACAATTTTGGGAGCAAGAGTCTCTCTTCCAGG 2370
QY 845 GluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuProAspArgSerAsp 864
DB 2371 GAAGATGGAGGCCAAGGAAATCTTATGTTAAATTTACTTCTTCCAGACAGAGTGAAT 2430
QY 865 LysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLysTrpAsnGlnThr 884
DB 2431 AAAACAAGAGAGAAACAAAACAGTCAAGAAAACCTTTGGAAACCAANTGGAACCACT 2490
QY 885 PheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluIleThrLeu 904
DB 2491 TTTCAATTTCTCTGTTTCCAGGAGAAATTTCCGGAAGAAATCTCGAATATCAACCTT 2550
QY 905 TrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGluIleLeuGlu 924

Db 2551 TGGATCAAGCTCGAGTTCGAGAGCAAGAAAGTGAATTCCTAGGAGAGATTTAATTGAA 2610
Qy LeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGlnThrHisAspVal 944
Db 2611 TTGGAACAGCTCTGTAGATGATGAACACACACACACACACACACACACACACACACACAC 2670
Qy 945 SerSerLeuProLeuProAspProSerProTyrLysLeuProArgArgGlnLeuHisGlyGlu 964
Db 2671 TCCTCATTTGCCACTCCCTCTCACCTTCTCCATATATGCGAGACACAGCTGCATGGAGAG 2730
Qy 965 SerProThrArgArgLeuGlnArgSerLysArgLysSerAspSerGluValSerAspTyr 984
Db 2731 AGCCCAACACGCGAGTTGCAAGGTGCAAGAGTAAGTGCACGGAAGTATCCGACTAT 2790
Qy 985 AspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGlyArgAspLeuGln 1004
Db 2791 GACTCGAGAGACGCGTGGAGTCTGTGCAGTATACCGACACAGTGGCGGTGATCTTCAA 2850
Qy 1005 SerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSerProSerGly 1024
Db 2851 AGCTCCACATATCAGTGGCAGAACAAAGTCATGTCATCAAAACCATGCTCACCATCAGGC 2910
Qy 1025 SerProHisArgValAspValIleGlyArgThrArgSerTrpSerProSerAlaProPro 1044
Db 2911 TCTCTCATCGAGTAGATGTTATAGGAAGCACTAGGTCACTGCTCGCTAGTGTCCCTCCT 2970
Qy 1045 ProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnThr 1064
Db 2971 CCTCAAGAGATGTGGAAACAGGGCTTCGAGGGACACGTGCTACTGGCCATTACATACA 3030
Qy 1065 IleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSerAspArgAsp--- 1083
Db 3031 ATTAGCCGGATGGACACACACCGTGTCTATGGACGACCATTAATCTTTCAGAGAGACAGT 3090
Qy 1083 ----- 1083
Db 3091 CATTTTCTCACTTACCTCGCTCTCGACACAGGCAGACAGTGAAGTCAACACAGGGAT 3150
Qy 1084 ---ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThrGluGln 1102
Db 3151 GGAAGGATTTGTGAAGCAGCAGATAGACAGCCATATCACAGATCCAGATCAACAGAAACA 3210
Qy 1103 ArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAspThrAsn 1122
Db 3211 CGGCTCTCTAGAGCGGACACACCCGCTCCAGATCCTTGAACGTGTGTATACAAAC 3270
Qy 1123 LeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAlaLeu 1142
Db 3271 CTATGAGGTCCGATGCCCTTCAATTAATGACTGGAAGATCTGCCCTCTTCACTGCCCTTA 3330
Qy 1143 SerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrProGlyThr 1162
Db 3331 TCGAGGTCTCACCTCGCATCGCTCTGTCTGACAGCCAGTCCGTCAGTACTCCCGTAACA 3390
Qy 1163 GlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg----- 1180
Db 3391 GGACGAAGGGCCCGCAGCTTCTCAGCTCCACCAAGGGAACCTGGAAAGATGATC 3450
Qy 1181 -----SerAlaMetAspIleGluGluArg 1188
Db 3451 ACAGAAGCATGGACTCCACAAGGAGAGAACTCAGGTGCTATGGATATAGAGAGAGA 3510
Qy 1189 AsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGlySerAspProArgLeuGlu 1208
Db 3511 AATGCCAAATGAACATTAACAAATACAAACACAGGTAGCCGATCAGATCCACAGACTGGAA 3570
Qy 1209 GlnAspTyrHisSerLysTyrArgSerGlyTrpAspProHisArgGlyAlaAspThrVal 1228
Db 3571 CAAGATTACCAATTGGAAGTATCGCTCAGGATGGGATCCACATAGAGGGGACAGATCTGTT 3630
Qy 1229 SerThrLysSerSerAspSerValSerAspValSerAlaValSerArgThrSerSer 1248
Db 3631 TCTACTAAATCCTCGGACAGTGAATGAAGTATGATCTATCTGCGGTTTCAGGACTAGTAGT 3690

Qy 1249 AlaSerArgPheSerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsn 1268
Db 3691 GCTTCTCGTTTCAGCAGCACAAGCTACATGTCGTTCATTCAGACCGCCGAGAGAAAC 3750
Qy 1269 ArgLysIleSerValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLys 1288
Db 3751 AGCAAAATCAGTCTCTTACGTCCAAAATGCAAAAGCAGACAGATGGTGTGTGCGGGAAG 3810
Qy 1289 AsnLeuThrLysSerThrSerIleSerGlyAspMetCysSerLeuGlnLysAsnAspGly 1308
Db 3811 AGCATGCCCAAAAGCACCAGCATCAGTGGAGACATGTGCTCTCACTGGAGAAATATGATGC 3870
Qy 1309 SerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyValLysArgSerSer 1328
Db 3871 AGCATGTCGACACTGCACTGGGCGCTCTGGGCAACGATGGCAAGAGCGCGCATCTAGC 3930
Qy 1329 IleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGln 1348
Db 3931 ATTGGGCGCAAAATGTTAGCTATTGTTGTTCTCTCACGGAAGTGCAGTGTCTCCAA 3990
Qy 1349 LeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerThrGlu 1368
Db 3991 CTCAGCCAAACCGAAGGAGAGTAAAGACTACGACGACTGTCCAGAGACACGAG 4050
Qy 1369 ThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThr 1388
Db 4051 ACCGGCTAGCAGTGGAGATGAGAACTGGATGACAGACAGCCAGCCGCGAATCCAGC 4110
Qy 1389 AspGlySerMetAsnSerTyrSerSerGluGlyAsnLeuIlePheProGlyValArgLeu 1408
Db 4111 GATGGCAGCATGAACATGATACAGCTCAGAGGAATCTGATCTTCTGGGGTCCGCTG 4170
Qy 1409 AlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGly 1428
Db 4171 GCCTCTCAGACCGAGTTTCTGATGCTGGGCTGGGCTGCTCAGCTAGTGGGA 4230
Qy 1429 ArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLys 1448
Db 4231 CGCCAGACCTCGCAGCTCTCTGCAATGGGTGACATTCAGGTGGGATGATGATGAAG 4290
Qy 1449 GlyGlnLeuGluValGluIleAlaArgGlyLeuValValLysProGlySerLys 1468
Db 4291 GGACAGCTGGAGGTAGATCATCCGGGCTCGAGGCTTGTAGTAAACCGGTTCAG 4350
Qy 1469 ThrLeuProAlaProTyrValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLys 1488
Db 4351 ACATGCGCAGCACAATATGTCAAGGTGTACTCTGTAGACAAACGAGTCTCATAGCCAAA 4410
Qy 1489 LysLysThrLysValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSerPhe 1508
Db 4411 AGAAACCAAGGTGGCGAAGACGCTGGAGCCCTGTATCAGCAGCTGTATCTTTTC 4470
Qy 1509 GluGluSerProGlnGlyArgValLeuGlnIleValTyrGlyAspTyrGlyArgMet 1528
Db 4471 GAAGAGACCCCGAGGAGGTGTACAGATCATTTGTCTGGGAGATNTAGCCGATG 4530
Qy 1529 AspHisLysSerPheMetGlyValAlaGlnIleLeuLeuAspGluLeuLeuSerAsn 1548
Db 4531 GATCACAAATCCCTTATGGAGTGGCCAGATACTCTTAGATCACTGGAATATCAAC 4590
Qy 1549 MetValIleGlyTrpPheLysLeuPheProProSerSerLeuValAspProThrSerAla 1568
Db 4591 ATGGTAATGGATGGTTCAAACTCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4650
Qy 1569 ProLeuThrArgArgAlaSerGlnSerSerLeuGluSerSerThrGlyProSerTyrSer 1588
Db 4651 CCTCTGACAGAGAGAGCTTCCCCAATCGTCTCTGGAAGTTCTACTGAGACCTTCTTACTCT 4710
Qy 1589 ArgSer 1590
Db 4711 CGTTCA 4716

```

RESULT 7
AF199324
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
1 (bases 1 to 5172)
Wang,Y., Sugita,S. and Sudhof,T.C.
The RIM/NIM family of neuronal C2 domain proteins. Interactions
with Rab3 and a new class of Src homology 3 domain proteins
J. Biol. Chem. 275 (26), 20033-20044 (2000)
20347919
MEDLINE
PUBMED
10748113
REFERENCE
2 (bases 1 to 5172)
Wang,Y. and Sudhof,T.C.
Direct Submission
Submitted (27-OCT-1999) Center for Basic Neuroscience, The
University of Texas Southwestern Medical Center, 6000 Harry Hines
Blvd., Dallas, TX 75335-9111, USA
FEATURES
source
Location/Qualifiers
1..5172
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..5172
/genes="Rim2-2B"
52..4251
/genes="Rim2-2B"
/codon_start=1
/product="RIM2-2B"
/protein_id="AAF81646.1"
/db_xref="GI:8925862"
/translation="MSAPLGPRGRPAPTAAOSPQPQPMPLDLSHTTEERKTIQAVM
DRQKEEKEQSVLKKLGHOFENYKEQVKWGEQSQEQEQKGDAPTCGICHTKPTAD
GCCHNGYCYOTKFCARGCGQSVLRKSNVWVNCMLCKRQBEILTKSGAWFNFSNTPQ
QDOKALGLRSEEAPOEKAKLHEOTQFGPGDSSVPAVERGRAHGLTRODSIKNG
SGMKHQIALSDPMDRKRSPVSRDQNRHIDQSEEREYEQYVPSDTPMRSPSDIADYR
SQRQEQFYEEPDLHNTVSRNGRHSHKEYIVDDDEVSREDVDEKQREESYQAIR
SDNLPAFYKPPQYEQMRIHAEVSRARHERHSDVLSANAELESPISLLEMDPRLS
RQSVASERRAAMQNSYMERTREAQGSQSPQRTNHSPTTPRSPISLDPELRL
ADSLRKQHLDPSSAVKTKRKQWETMLRNDLSLSDQSVSRVPPRPPIKSKGKGWR
QVLSRSHLESLASTPVTSCDDVYIESESVEKGDQKGRKTSQCAVWSDSNTRSER
QKVMYSGHSLDLEDLWSEPTQKDSGVTCSSTLINEERSHSDHEPVTWQPSKQDR
LIGRIILNKLKGVSRSDGMLGKLVGGKWTSGRGLCAFITKVKKGLSLATVAGHLE
RPGDEVLMNKRLLQATFEENVNILESKPEQVELVWSRPIGDMPRIPDSVTHAQL
SSSSSESQMDRPSIVTSVPMGPKMDRVPQFLSQGLSIKLMFDRVGHQLITVLGA
KDLPSDEGRPNRPYKIVFLPDRSDKNKRKTVKTKLEPKMNQTFIVSPVHRREFR
ERMLETLMDQARPEREESFEGLIELLTLALDDDEPHWYKLTQHDVSLPLPSPSP
VMPRCLHGESPTRLQRKRIISDSVSDYDCEDGVGVSDVHEDGRDLQSSLTLYPE
QVMSNHCSPSGSPHVDVIGTRTSWSPVPPQNRVQGLGTRATGHTYNTLSMRDR
HRVMDTHYSERESHFTLPRSHRQTSSEHHDRGRDCEAADQPVHRSRSTSRQRLP
EYTTTSSRSEADATNLMRSMPSLMTGRSAPPALSRSHPTGSGVTSPTSPSTVTVGR
RGSGLPOLPFGKTLRGGGKKLSTVQRSTETGLAVEMRNMTWQASRESTDGSMSY
SSNGENILFPQVTLASQFSDFLDGLGPAOLVGRQTLTAPMGDIOVGMDKKGLEL
EILRARGLVVPGSKTLPAPYKVKVYLLDNGVCIATKKTIVAKTLEPLVQQLLSFRES
PQKVLIQITVQVIGRMDKSPFMGVAQILLDELELSNMVIGNFKLFPFSLVDPTLAL
LTPRASQSSLSSTGSGYSRS"
1555 a 1190 c 1286 g 1141 t
BASE COUNT
ORIGIN

```

Alignment Scores:	
Pred. No.:	2,44e-207
Score:	6856.50
Percent Similarity:	84.06%
Best Local Similarity:	83.00%
Query Match:	82.76%
DB:	10
	3
	235
	Indels:
	27
	Mismatches:
	12
	Conservative:
	1338
	Matches:
	5172

	Qy	Db
341	HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr	360
341	HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr	360
892	GAAACCTCAATTTTATGAAGAACCTGATCATTTAAATTTATAGGAVTCTTAACAGGAGAGCC	951
321	GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgGly	340
832	CCCTCAGATAGACAAATGCTAGATCTCATCAGATTATGCTGATAGACAGATCTCAACG	899

Db 952 CATAGACATTCCAAAGAGTATATTTGTAGACGACGAAGATGTGAGAGCAGAGATGAATAT 1011
Qy 361 GluArgGlnArgArgGluGluGluGluGlnAlaArgTyrArgSerAspProAsnLeuAla 380
Db 1012 GAAGAACAAGAGAGAGAGAGATACCGAGCAGCTACAGAAAGTATCCAAATTTGGCC 1071
Qy 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
Db 1072 CGGTATCCGGTAAAGCCACCAACCCCTATGAAGAGCAAAATGCGGATCCACGCTGAAGTGTCC 1131
Qy 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAlaGluLeuGluAsp 420
Db 1132 CGGCACGACACAGAGAGAGCAGAGTATGTTCTTTGGCAAAATGCTGAACTAGAAGAT 1191
Qy 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
Db 1192 TCCAGGATTTCTCTGCTAGAGATGATAGACCATCAAGCGAAAGATCTGTATCTGAAGCT 1251
Qy 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
Db 1252 AGAGCTGCAATGGAAACACGAGATCGTATTCAATGGAAAGAACTCGAGAGCTCAGGGA 1311
Qy 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
Db 1312 CAAAGATTCTTATCCCAAAGAGCACCAATCATAGTCTCTCTACCCCTCGCAGAGCCCT 1371
Qy 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisIleu 500
Db 1372 ATACCGCTCGATAGACAGAGCTGAGCGTGCCAGCTCCCTACCGAAACAAACACACTTA 1431
Qy 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
Db 1432 GATCCAGCTCTGCTGTAAAGAAACGAGAGCGAGAAATGGAAACAACTGTGAGGAAT 1491
Qy 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
Db 1492 GATTCTTTGAGTTCAGACCACTCTGAGTCAGTGAGCGCGCCGCCACCAAGGCCCTCATAAA 1551
Qy 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAla 560
Db 1552 TCCAAAGAAAGAGGTAAATCGCCAGGTTTCATCTGAGCAGCTCAGAGAGAGGTGGCA 1611
Qy 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
Db 1612 TCCACGCTGAGTATACAGCTGTGATGACGTGAGATTGAAGCGAGAGGTAGGTGAG 1671
Qy 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1672 AAAGGACACAGTCAAAAGGGGAAAGAAAGAACTAGTGAGCAGGCGAGTTTGGTCGAGCTCT 1731
Qy 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1732 AACACCAAGGCTGAGAGACAAAGAAAGATGATGATCTCTGTGTGGCCACTCTTTTGGACGAG 1791
Qy 621 AspLeuGluTyrSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
Db 1792 GATTTGGATGGTCTGAGCCTCAGACTAGAGACTCTGSGGTAGACACCTGTAGTAGCACA 1851
Qy 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrThrGlnProSerLys 660
Db 1852 ACTCTTAACGAGGAGCGTAGCCATAGTGATGAGCACCCTGTGACCTGGACGCCATCCAAA 1911
Qy 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 1912 GATGAGATCCGCTAAATTTGGTCTGATTTATTTATTAATAGCGTTTAAAGATGGGAGTGTA 1971
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 1972 CCTCGAGATTGAGAGCAATGCTGGGCTTAAGGTTTGAAGGAGGAAGATGACTGAATCA 2031
Qy 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 2032 GGTGCGATTTGTGCAATTTATTAACCAAGTGAAGAAAGGAAGTTTAGCTGACTGTAGGA 2091

Qy 721 HisLeuArgProGlyAspGluValLeuGluTyrAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 2092 CAICTTAGACCAAGGTGATGAGTCTTGGATGGATGGAGGCTATTGCGAGGAGCCACA 2151
Qy 741 PheGluGluValTyrAsnIleLeuLeuGluSerLysProGluProGlnValGluLeuVal 760
Db 2152 TTTGAGGAAGTTTACCAATATTATTAGATCCCAAGCTGAAACCAAGTTGAGCTGTGT 2211
Qy 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2212 GTTTCAAGGGCCGATTGAGATATGCCTAGAATACCTGATAGCACTCATGCAACACTGAAA 2271
Qy 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2272 TCCAGTTCTAGCTCATTTGATCTCAAAAATGGATGCTCGTCTATATGCTGTACTCT 2331
Qy 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2332 CCCATGAGTCTGGCATGCTGAGGAGTGTCCACAGTTCTTATCTGACAGCTTTCATA 2391
Qy 821 LysLeuTyrPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2392 AAACATATGTTGATGACAGGTGTGTCACAGTTAATAGTTACAATTTTGGGAGCAAGAT 2451
Qy 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2452 CTCCCTTCCAGGAGATGGAGGCCGAGGATCTTATGTTTAAATTTACTTCTTCCCA 2511
Qy 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2512 GACAGAGTGTATAAAACAAGAGAGAAACAACAACAGTCAAGAAACTTTTGGACCCAAA 2571
Qy 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2572 TGGAACACAGACTTTCATTATTCTCTGTTTCCCGAAGAGATTTCCGGAGCAAGATGCTC 2631
Qy 901 GluIleThrLeuTyrAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920
Db 2632 GAATACACCTTTGGGATCAGCTCGAGTTCGAGAGGAGAAAGTGAATTTCTTAGGAGAG 2691
Qy 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspGluProHisIleTyrLysLeuGln 940
Db 2692 ATTTTAATGATGGAAACAGCTCTGTAGATGATGAACAACACACTGTGTACAACTCAG 2751
Qy 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
Db 2752 ACCCATGATGCTCTCTCATTTGCCACTCCCTCACCTTCTCCATATATGCCCAGGAGAG 2811
Qy 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db 2812 CTGCATGGAGAGAGCCCAACACAGCGGTTCGAAGGTTCGAAGAGAAATAGGTGACAGCGAA 2871
Qy 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
Db 2872 GTATCCGACTATGACTCGAGAGCGGTGGAGTCTGTCTCAGATTACCCAGACAGATGGC 2931
Qy 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
Db 2932 CGTGATCTTCAAGACTCCCATTTATGATGTCGAGCAAGCTCATGTCTATCAACCAATGC 2991
Qy 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
Db 2992 TCACCATCAGGCTCTCTCTCATCGATAGATGTTATAGAGAGGACTAGGTCTATGTCGCT 3051
Qy 1041 SerAlaProProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 3052 AGTGTCTCTCTCTCTCAAGGAATGTGAACAGGGCTTCGAGGACACGCTGTACTTGGC 3111
Qy 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 3112 CATTAACAATACAAATTAGCCGATGGACAGACACCGTGTCTATGACGAGCACTTACTCTTCA 3171

QY 1081 AspArgAsp----- 1083
Db 3172 GAGAGAGAGCAGTCATTTTCTCACTTACCTCCCTCTCGACACAGCAGCAGCAGTGGAGCAT 3231
QY 1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
Db 3232 CACCACAGGAGTGAAGGAGTGTGAAGCAGCAGATAGACAGCCCATATACAGATCCAGA 3291
QY 1099 SerThrGluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArg 1118
Db 3292 TCACAGACACACAGCGCTCTCTAGAGCGGACACACCCGCTCCAGATCCCTCTGAACGT 3351
QY 1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
Db 3352 GCTGATACAACTCATGAGTTCGATGCTTCATTAAATGACTGGAGATCTGCCCTCTCT 3411
QY 1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
Db 3412 TCACCTGCTTATTCGAGGCTCTCACCTCGCACTGGCTCTGTCCAGACAGTCCGTCAAGT 3471
QY 1159 ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeu 1178
Db 3472 ACTCCGTTACAGGACGAGGGCGCGAGCTTCCTCAGCTCCCAACCAAGGGAACACTG 3531
QY 1179 GluArgSerAlaMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrLys 1198
Db 3532 GAAAGA----- 3537
QY 1199 GlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGly 1218
Db 3537 ----- 3537
QY 1219 TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSer 1238
Db 3537 ----- 3537
QY 1239 AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMet 1258
Db 3537 ----- 3537
QY 1259 SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet 1278
Db 3537 ----- 3537
QY 1279 GlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGly 1298
Db 3537 ----- 3537
QY 1299 AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu 1318
Db 3537 ----- 3537
QY 1319 GlyThrSerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGly 1338
Db 3537 ----- 3537
QY 1339 LeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLys 1358
Db 3538 -----CGAGCAGGTAAAAAG 3552
QY 1359 LeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrp 1378
Db 3553 CTACGAGCAGCTGCTCCAGAGAGCAGCAGCGGCTAGCAGTGGAGATGAGGACTGG 3612
QY 1379 MetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGlu 1398
Db 3613 ATGACAAGACAGCCAGCGGGAATCCAGGATGGCAGCATGAACATGTACAGCTCAGAA 3672
QY 1399 GlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeu 1418
Db 3673 GGAAATCTCATCTTCTCGGGTCCGCTCTGACAGCCAGTTCAGCGATTTCCCT 3732
QY 1419 AspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly 1438

Db 3733 GATGTCCTGGCCCTGCTCAGCTAGTGGAGCGCCAGACCCCTGGGACTCTCTGCAATGGGT 3792
QY 1439 AspIleGlnValGlyMetMetAspLysGlyGlnLeuGluValGluIleIleArgAla 1458
Db 3793 GACATTCAGTGGGGATGATGATGAAGAGGACAGCTGGAGGTAGAGATCATCCGGGCT 3852
QY 1459 ArgGlyLeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyr 1478
Db 3853 CGAGCCCTTGTAGTAAACACAGGTTCCAGACACTGCCAGCACCATATGTCAAGTGTAC 3912
QY 1479 LeuLeuAspAsnGlyValCysIleAlaLysLysLysThrLysValAlaAArgLysThrLeu 1498
Db 3913 CTGTTAGCAACGAGTCTGATAGCCAAAGAAACCAAGTGGGAGAGAGCGCTG 3972
QY 1499 GluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGln 1518
Db 3973 GAGCCCTTGTATCAGCAGCTGTTATCTTTCAGAGAGAGCCCCAGGGAGAGTGTACAG 4032
QY 1519 IleIleValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGln 1538
Db 4033 ATCATTCCTGGGAGATTAAGCGCATGATCACAATCTTTATGGAGTGGCCCG 4092
QY 1539 IleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPhePro 1558
Db 4093 ATACTCTTAGATGAATGGAATATCCAACTGATGATGATGATGATGATGATGATGATGAT 4152
QY 1559 ProSerSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSer 1578
Db 4153 CCTCTCTCTCTAGTAGATCCAACTTGGACCTCTGACAGAGAGCTTCCCAATCGTCT 4212
QY 1579 LeuGluSerSerThrGlyProSerTyrSerArgSer 1590
Db 4213 CTGGAAGATTCATCTGACCTTCTTACTCTCTGTTCA 4248

RESULT 8
AF199329 4965 bp mRNA linear ROD 04-JUL-2000
LOCUS Rattus norvegicus RIM2-4C (Rim2-4C) mRNA, complete cds.
DEFINITION AF199329
ACCESSION AF199329
VERSION AF199329.1 GI:8925871
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 4965)
AUTHORS Wang, Y., Sugita, S. and Sudhof, T.C.
TITLE The RIM/NIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins
J. Biol. Chem. 275 (26), 20033-20044 (2000)
JOURNAL 20347919
MEDLINE 10748113
PUBMED
REFERENCE 2 (bases 1 to 4965)
AUTHORS Wang, Y. and Sudhof, T.C.
TITLE Direct Submission
JOURNAL Submitted (21-OCT-1999) Center for Basic Neuroscience, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, TX 75235-9111, USA

FEATURES
source
1. 4965
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
gene
1. 4965
/gene="Rim2-4C"
52..4044
/gene="Rim2-4C"
/codon_start=1
/product="RIM2-4C"
/protein_id="AAF81651.1"
/db_xref="GI:8925872"

/translation="MSAPLGRGRPAPTPAASPPPPQPPMPDLSHLTEERKTIQAVM
 DRQKEERQSVLKHLQOPEMYKQVKKQGESEQQQEQKGDAPTGGICHKTKFAD
 GCCHNGSCYCTQCARCGRSVLSRKNKVMVCMKQKQVPAVERGRAGHGTRODSIKNG
 QPDKALRGSRSEAPQKAKLHEOTOFOGPPGDSVPVAVRGRAGHGTRODSIKNG
 SGNKHQIASMPDRKSPSVSDQNRERVDQSEEREYSQYVPSDSTMPSRSPDYADR
 RSOREFOYIEPDLNTRNRRHRSKSEYIVDDSEVESRDEYERQREESTQARIR
 SDPNLARYPKQPIYEOEMRIHAVERARHRRHSDVSLANAELESDRSISLMDRPS
 RORSVSERRAAMNORSYSMERTREAGQSSYPQRTTNHSPPTPRSPISLDPELR
 ADSLRQHLDPSSAVRKTREKMTWLRNDSLSQDSSESVRPPRPHKSKGGMWR
 QVLSSESEELAPTEVTSDDVEIESVSGEKDMEYSLWLEHSHSSSEASPMSLPH
 VTWQPSKGDGELLGRILLNKLKDGVSPPDSGAMGLKVVGGKMTSSEGLCAITKVK
 KGLADTVGHLRPGDEVLWNGLLQOAPFEVINILSKPBPQVELVVSRLDIMP
 RLPDTHAQLESSESSSESQSDPSITSVSPMSPGMLRDVPOFLSQSLKLMFQVY
 GHQIVITLGAKDLPSRSDGRPNPYVYKIFLDPDRDKNRRTKTKVKTLEPKWNTF
 IYGVPRHREFREMLTTLMDQARVEESEFLGEILLLELTALLDDEPHYKQTHD
 VSSLPLPHSPMPRLHGESPTRLQSRKISDSRVSYDCEDXGVGVSDYRHDGR
 DLQSSLTSLVPEQVMSNHCSPPSPRVDVIGTRSWSPSPVPPQNRQEGLAGTAT
 GHTNLSMRDRHVMDDHSEDRDCEAADROPYHRSRSTEQRLPHERITTRSRSE
 RADTNLMSRMSPLMGTSRAPSPALSRSHPTGVSOTSPSTPVLURGRQLPOLPPK
 GTLERGGKKLRSTVQSTETGLAVENRMNTQASRESTDGMNYSSEGNLIFPGV
 RLASDSQFDFDLGLGPAQLVGRQTLATPANGDIQVGMMDKQGLVBIIRARGLVVK
 PGSKTLPAPVYKVVLLDNGVCIAKKTIVARKTLEPLVQQLSFEPSQGVKLIIVM
 GDYGRMDHKSFMVQAIIIDLELSNVICWFKLFFPSSLDVTLAPLTRASQSSIE
 SSTGPSYSYRS"

BASE COUNT 1486 a 1145 c 1232 g 1102 t
 ORIGIN

Alignment Scores:

Pred. No.: 1,67e-197 Length: 4965
 Score: 6543.00 Matches: 1283
 Percent Similarity: 81.82% Conservative: 18
 Best Local Similarity: 80.69% Mismatches: 29
 Query Match: 78.97% Indels: 260
 DB: 10 Gaps: 4

US-09-617-099B-1 (1-1590) x AF199329 (1-4965)

Qy 1 MetSerAlaProLeuGluGlyProArgGlyValArgProAlaProThrProAlaAlaSerGlnPro 20
 Db 52 ATGTCGGCTCCACTCGGGCCCCGGGGCGCGCGGCTCCACCCCGGGCGCTCTCAGCCT 111
 Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysLeile 40
 Db 112 CTTCCGAGCCGAGATGCCGACCTCAGCCACCTCAGGAGAGAGAGAGAGAGAAATCATC 171
 Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysleile 60
 Db 172 CAGGCTGTCTATGGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
 Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
 Db 228 ----- 228
 Qy 81 ValAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
 Db 229 -----AAG 231
 Qy 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
 Db 232 CTGCATCAACAAATTTGAATGTATAGGAGCAAGTCAAGAGAAATGGGAGAGAGATCAAA 291
 Qy 121 GlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
 Db 292 CAGCAGCAGAT 351
 Qy 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
 Db 352 GCAGATGGATCGGCCCACTACTGTTCATATGCCAACCAGATTCCTGTGCTGTGTGGA 411
 Qy 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
 Db 412 GGTGCGATGTCTATACGCTCAACCAAGATTATGTGGGTGTGTAATTTGTGCCGAGAAACAA 471

181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
 472 CAAGAAATCTCTACTAAGTCGGCGCGTGTGTTTACATAGTGGTCTTAATACACCGCAG 531
 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGlnLysLys 220
 532 CAGCTCGATCAAAAGGCTCTTCGAGGGCTTCGCGAGTGAAGAGAGAGAGAGAGAGAG 591
 221 AlaLysLeuHisGlnGlnProGlnPheGlnGlnValAlaProGlyAspLeuSerValProAla 240
 592 GCAAAATCTCATGAGCAGCAGCGAGTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrLysLysAsnGlySerGly 260
 652 GTTGAGAGAGAGTCTCAGTCTCATGGCTCAACAAGCAGAGATTCATTATTAAGATGCTCAGGA 711
 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 712 ATGAAGCAGCAGATCCCGCAGTGCATGCCCTTCACAGAGAGAGAGAGAGAGAGAGAGAG 771
 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTrpVal 300
 772 AGGGATCAAAATCGAAGATACGACCAAGTGAAGAGAGAGAGAGAGAGATTCACAGTATGTT 831
 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
 832 CTTTCAGATAGCAATGCTAGATCTCCATCAGATTATGCTGATAGAGATCTCAACGT 891
 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgLys 340
 892 GAACCTCAATTTTATGAAGAACCTGATCATTTAAATATATAGGATTCCTAACAGAGAGGC 951
 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
 952 CATGACATTCCTCAAGAGATATATTGAGACGAGAGATGTCGAGAGAGAGATGAATAT 1011
 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
 1012 GAAAGACAAAGAGAGAGAGAGAGAGAGTACCAGGACCGCTACAGAGTGTATCCAAATTTGGCC 1071
 381 ArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSer 400
 1072 CGGTATCCGCTAAAGCCACAAACCTATGAAGAGCAAAATGCGATCCAGCTCAGAGTGTCC 1131
 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
 1132 CGGCGACGACACAGAGAGAGAGAGAGAGTGTATGTTCTTTGGCAAAATGCTGAACATAGAAT 1191
 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
 1192 TCCAGAGATTTCTGCTTAGGATGGATAGACCATCAAGGCAAGAGATCTGTATCTGACGT 1251
 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
 1252 AGAGCTGCAATGAGAAACCCAGCGATCGTATTCATGAAAGAACTCGAGAGAGTTCAGGGA 1311
 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
 1312 CAAAGTTCCTTATCCAAAGAGAGAGAGAGATCATAGTCTCTCTACCTCCCGAGGAGCCT 1371
 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLysLeu 500
 1372 ATACCGCTCGATAGACAGAGCTGAGCGGTGCCCATCTCCCTACGAAACACACACCATTA 1431
 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
 1432 GATCCCGAGCTCTGCTGTAGGAGAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1491
 521 AspSerLeuSerSerSerAspGlnSerGluSerValArgProProProProProProHisLys 540
 1492 GATTCTTTGAGTTGAGACAGAGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1551
 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560

Db 1552 TCCAGAAAGGAGGTAATAATGCCAGAGTTTCACTGACAGCTCAGAGAGGAGTTGCCA 1611
Qy SerThrProGluTyrThrSerCysAspValGluLeuGluSerGluSerValSerGlu 580
Db 1612 TCCAGCCTGAGTATACAAAGCTGTGATGACGTGGAGATTGAAGACGAGAGGTAGTGAG 1671
Qy LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1672 AAAGAGACATGGAG----- 1686
Qy 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGlu 620
Db 1687 -----TACAGCTGGTTGGAG 1701
Qy 621 AspLeuGluTyrSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 640
Db 1702 CATCGCTCTGGCAT----- 1716
Qy 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTyrGlnProSerLys 660
Db 1717 -----AGCAGTGGGCTATCCCAATGCTTTTGCACCTGTGACCTGGCAGCCATCCAAA 1770
Qy 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 1771 GATGAGATCGCTAATTTGGTCTATTTTATTAATTAAGCTTTTAAAGATGGAGTGA 1830
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 1831 CCTCGAGATTCAGGAGCAATGCTGGCTTAAAGTTTGAAGGAAAGTACTCAATCA 1890
Qy 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 1891 GGTGACATTTGTGCATTTATTAACCAAGTGAAGAAAGGAGTTTGTAGTACTGTAGGA 1950
Qy 721 HisLeuArgProGlyAspGluValLeuGluTyrAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 1951 CATCTTAGACAGGTGATGAAGTCTTGGATGGATGGAGCTATTTCAGAGGACCCACA 2010
Qy 741 PheGluGluValTyrAsnIleLeuLeuGluSerLysProGluProGlnValGluLeuVal 760
Db 2011 TTTGAGGAAGTTTACAACATTATTCTAGAACTCAAGCCTGAACCAACAAAGTTGAGTGT 2070
Qy 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2071 GTTTCAGGCCCATTTGGAGATATGCTAGAAATACCTGATAGCACTCATGCACAACATG 2130
Qy 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2131 TCCAGTTCTAGTCTCAATTTGAATCTCAAAAATGGATCGTCCGTCTATATCGGTACCTCT 2190
Qy 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2191 CCCATGAGTCTGGCATGCTGAGGAGATGCCACAGTCTTATCTGCAAGCTTTCAATA 2250
Qy 821 LysLeuTyrPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2251 AAACATGTTTGCACAGGTTGCTCACCAGTTAATAGTTTACAAATTTTGGAGCAAGGAT 2310
Qy 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2311 CTCCCTTCAGGAGATGGAGGCCAAGAAATCTTAATGTTTAAATTTACTTCTTCCA 2370
Qy 861 AspArgSerAspLysAsnLysArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2371 GACAGAGTGATAAAACAAGAGAGAGACAAAAACAGTCAAGAAACCTTTGGAAACCCAAA 2430
Qy 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2431 TGGAACACAGACTTCAATTTATTTCTCTGTTCCCGAAGAGAAATTCGGGAACGAATGCTC 2490
Qy 901 GluIleThrLeuTyrAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920

Db 2491 GAAATCACCCCTTTGGGATCAAGCTCGAGTTCGACGAGGAGAAAGTGAATTCCTAGGAGAG 2550
Qy 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisIleTyrTyrLysLeuGln 940
Db 2551 ATTTTAATTAATTTGGAAACAGCTCTCTGTAGATGATGAACACACACTGGTCAAACTTCAG 2610
Qy 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgGln 960
Db 2611 ACCCATGATGCTCTCTCATTTGCCACTCCTCACCTTCTCCATATATGCCACGAGACAG 2670
Qy 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db 2671 CTGCATGAGAGAGCCCAACACAGCAGGTTCGAAAGGTCGAAAGATTAAGTGCAGCGAA 2730
Qy 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
Db 2731 GTATCCCACTATGATCGAGGACGCGGTGGAGTCTGTGTCAGATTATCCGACACGATGCG 2790
Qy 1001 ArgAspLeuGlnSerSerThrLeuSerValProGlnGlnValMetSerSerAsnHisCys 1020
Db 2791 CGTGATCTTCAAGCTTCAATATCATGTCGACAGACAGTCAATGTCATCAACCATTCGC 2850
Qy 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTyrSerPro 1040
Db 2851 TCACCATCAGCTCTCTCATCGAGTAGATGTTATAGGAAGACTAGTCTCATGCTCGCCT 2910
Qy 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 2911 AGTGTCCCTCTCTCTCAAGGAATGTGGAACAGGGGCTTCGAGGACACGCTGCTACTGCG 2970
Qy 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 2971 CATTCACATACATTAAGCCGATGACAGACACACCGTGTGTCAGACGACCATTACTCTCA 3030
Qy 1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
Db 3031 GAGAGAGACAGGATTTGTAAGACGACAGATAGACAGCCATATCAGATCCAGATCAACA 3090
Qy 1101 GluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAsp 1120
Db 3091 GAACAACGGCTCTCTAGAGCGGACCAACCGCTCCAGATCTCTGAAACGTGCTGAT 3150
Qy 1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
Db 3151 ACAACCTCATGAGTCGATGCCCTTCAATTAAGTGTGAAGATCTGCCCTTCTCTCACCT 3210
Qy 1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
Db 3211 GCCTTATCGAGTCTCACCTCGCACTGGCTCTGTCCAGACCACTCCGTCAAGTACTCCC 3270
Qy 1161 GlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
Db 3271 GTAAACAGACAAAGGGCGGCGAGCTTCTCAGCTCCCAACAAAGGGAACTGTGAAGA 3330
Qy 1181 SerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
Db 3330 ----- 3330
Qy 1201 AlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTyrAsp 1220
Db 3330 ----- 3330
Qy 1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspVal 1240
Db 3330 ----- 3330
Qy 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerVal 1260
Db 3330 ----- 3330
Qy 1261 GlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsn 1280
Db 3330 ----- 3330

Qy 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMet 1300
 Db 3330 ----- 3330

Qy 1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320
 Db 3330 ----- 3330

Qy 1321 SerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
 Db 3330 ----- 3330

Qy 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyGlyLysLeuArg 1360
 Db 3331 ----- 3331

Qy 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
 Db 3352 AGCACTGTCCAGAGAAGCAGCGAGACCGGCTAGCAGTGGAGATCAGCACTGGATGACA 3411

Qy 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn 1400
 Db 3412 AGACAGGCCACCGCGGAATCCAGCATGCGAGCATGAACAGTTACAGCTCAGAAGGAAT 3471

Qy 1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
 Db 3472 CTGATCTTCTGGGTCCGCTGGCTCTGACAGCCAGTTACAGCATTTCTTGATGT 3531

Qy 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
 Db 3532 CTGGCCCTGTCTAGTGGAGCCGACAGCCCTGGGCACTCTCTGCAATGGGTGACATT 3591

Qy 1441 GlnValGlyMetMetAspLysGlyGlnLeuGluValGluIleLeuArgAlaArgGly 1460
 Db 3592 CAGTGGGGATGATGATGAAGAAGCAGCTGGAGGTAGAGATCATCCGGGCTCGAGGC 3651

Qy 1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu 1480
 Db 3652 CTGTGTAGTAAACAGGTTCCAGACATCTGCAGCACCATATGTCAAGGTGATCTGTTA 3711

Qy 1481 AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaAlaArgLysThrLeuGluPro 1500
 Db 3712 GACACAGGAGTCTGCATAGCCAAAGAAACCAAGGTGGCGAGAAAGACGCTGGAGCCC 3771

Qy 1501 LeuTyrGlnGlnLeuLeuSerPheGluGlnSerProGlnGlyArgValLeuGlnIleIle 1520
 Db 3772 CTGTATCAGCAGCTGTATCTTTCGAGAGAGAGCCCGGAGAGGTGTTCAGATCAT 3831

Qy 1521 ValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
 Db 3832 GTCTGGGGAGATTATGGCCGATGATGATCAAAATCTTTATGGAGTGGCCAGATATCTC 3891

Qy 1541 LeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPheProSer 1560
 Db 3892 TTAGATGAATGGAACTATCAACATGATGATGATGATGATGATGATGATGATGATGAT 3951

Qy 1561 SerLeuValAspProThrSerAlaProLeuThrArgAlaSerGlnSerSerLeuGlu 1580
 Db 3952 TCCCTAGTAGATTCACAACTTGGACCTCTGACAGAGAGAGTTCCTCAATCGTCTCTGAA 4011

Qy 1581 SerSerThrGlyProSerTyrSerArgSer 1590
 Db 4012 AGTTCTACTGGACCTCTTACTCTCTGTTCA 4041

RESULT 9

AF199323
 LOCUS AF199323 5031 bp mRNA linear ROD 04-JUL-2000
 DEFINITION Rattus norvegicus RIM2-2A (Rim2-2A) mRNA, complete cds.
 ACCESSION AF199323
 VERSION AF199323.1 GI:8925859
 KEYWORDS Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

1 (bases 1 to 5031)
 Wang, Y., Sugita, S. and Sudhof, T.C.
 The RIM/MIM family of neuronal C2 domain proteins. Interactions
 with Rab3 and a new class of Src homology 3 domain proteins
 J. Biol. Chem. 275 (26), 20033-20044 (2000)

JOURNAL

20347919

MEDLINE

10748113

PUBMED

REFERENCE

2 (bases 1 to 5031)

Wang, Y. and Sudhof, T.C.

Direct Submission

Submitted (27-OCT-1999)

Center for Basic Neuroscience, The

University of Texas Southwestern Medical Center, 6000 Harry Hines

Blvd., Dallas, TX 75235-9111, USA

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

J. Biol. Chem. 275 (26), 20033-20044 (2000)

FEATURES

source

1. 5031

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

1. 5031

/gene="Rim2-2A"

52. 4110

/gene="Rim2-2A"

/codon_start=1

/product="RIM2-2A"

/protein_id="AAF81645.1"

/db_xref="GI:8925860"

/translation="MSAPLGRGRPAPTAASQPPQPEMPDLHSLTEERKIIQAVM

DROKREKESQVLKCLHQFQYKQVKGESQOQKQDGAFTGCHTKTKPAD

CGKHCSCYCTKFCARCGSVLSRNVKVCNLCRQKQILTKSGAFVNSGNTQ

QPKALGRLSEARQEKALHEQTFQFGPDSSVPAVERAGHLTRQDSIKNG

SMKHQIASDMPDRKSPSDQNRKYDQSRSEIYQTVPSDSTWPSPPSYAR

RSRQEPQFPEPLNYRDSNRHRSKTYTVDEDSRDEYERRESEYQARY

SDNLARYPKQPYEQRIHAESRAREHRSYVLSANAELEDRIILMDRPS

ROBSRRRAAMENORSYMERTRAQSSYPORTNHSPTPRRPIPLDRPELR

ADSLRKQHLDPSSAVKTKKEMETMLRNDLSLSDQSESVPRPPPHSKKQKGR

QVLSSESELASTPEVTSDDVIESVGEKDMYSLWLEHASSHSSASPSLHP

VTQPSKQDRLILGLINKLKDQSVPRSGAMGLKLVKVGKNTSEGLCAFTTKVK

KGLADTVGLHRLFGDEVLWNLQGLQATFEVYNILLESKPEQVELVSRPTGMP

RIPTDTHAQLESSESSPSQMDRPSISVTPSPMSGMLRDPQFQSLQSLKLMFVK

GHOLIIVILAKOLPSREDGRPNPYKIVFLPDRSKKRTNTVKTLEPKWQTF

IYVPRFERFERMLEITLMDQARVERESEFELEILETALDDPEHMYKLTND

VSLPLPHSPYPRRQLHGESPTRLQSKRISDSVSDYDCDGVGVSDYKDR

DLQSLUSVEQVMSNHSPSGPHRVIVIGTRSWSPVPPQPVNVEGLRTRAT

KRSRTEQRLILLRTTSSRSEADTNLMSPLMTGRSAPSPALSRSHPTGVSQ

TSPSTPTVGRGRQLPOLPKGLTBRGGKCLRSVTORSTETCLAVEMNMTQAS

RESTDGSMNYSSEGNLIFFQVLASDQSPDLGLGPAQLVCRQLATPAMGDIQV

GMDDKQQLSEIIRARGLVVPGSKTLPAPYKVLDDNGVCLAKKTKVARTKLP

LYQQLSFEPSQGVQLIIVMGDYGMDHKSFMGVAQIILDLLELSNMVIGWPKLFP

PSSIVDPTLAPLTRASQSSLESSTGPSYRS"

BASE COUNT 1503 a 1166 c 1246 g 1116 t

ORIGIN

Alignment Scores:

Pred. No.: 7,72e-197 Length: 5031

Score: 6522.00 Matches: 1283

Percent Similarity: 80.71% Conservativeness: 18

Best Local Similarity: 79.59% Mismatches: 29

Query Match: 78.72% Indels: 282

DB: 10 Gaps: 5

US-09-617-099B-1 (1-1590) x AF199323 (1-5031)

Qy 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20

Db 52 ATGTCGGCTCACTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTCACGCT 111

Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIleLeu 40

Db 112 CCTCGCAGCCGAGATGCCGACCTCAGCCACCTCAGCGAAGAGGAGGAAAATCATC 171
Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysIle 60
Db 172 CAGGCTGTGTCATGATCGTTCAGAGAAAGAGAGGAGGACGTCGTGCTCAAA--- 228
Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 228 ----- 228
Qy 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
Db 229 -----AAG 231
Qy 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGlnSerGln 120
Db 232 CTGCATCAACAATTTGAATGATATAGGAGCAAGTCAAGAAATGGGAGGAAATCACAA 291
Qy 121 GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
Db 292 CAGCAGCAGAGCAGAGAGGCGCAGCCCGACCTCGGCATCTGCCACAGACAAATTT 351
Qy 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
Db 352 GCAGATGATGGCGGCATAACTGTTCAATATGCCAAACCAAGTTCGTGCTGTTGGA 411
Qy 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
Db 412 GGTGAGTGTCTAATACGCTCAACCAAGGTTATGTGGTGTGTAATTTGTGCCGAAACAA 471
Qy 181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
Db 472 CAAGAAATCTCTCAATAGTCGGCGCGTGTGTTTACATAGTGGTCTAATACACCGAG 531
Qy 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLysLys 220
Db 532 CAGCCTGATCAAAAGGCTCTCGAGGCTTCGAGTGGAGAGGCCCTCAGGAGAGAG 591
Qy 221 AlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
Db 592 GCANAACTGCATGAGCAGACGAGTTTCCAGGACCCCGGTGACTCATCATCTGCA 651
Qy 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
Db 652 GTTGAGAGAGGTCGAGCTCATGGCTCACAGACAGATTCATTAAAGATGGCTCAGA 711
Qy 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
Db 712 ATGAAGCACCATGCCAGTCAGATGCCCTTCAGACAGAAAGAAAGTCCGTCACTGTC 771
Qy 281 ArgAspGlnAsnArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
Db 772 AGGGATCAAAATCGAAGATACACCAAGTGAAGAGAGAGGAAATATTACAGTATGTT 831
Qy 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
Db 832 CTTTCAGATAGCACATGCTAGATCTCCATCAGATTATGCTGATAGACATCTCAAGT 891
Qy 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
Db 892 GAACCTCAATTTATGAAGAACTGATCATTTAAATATAGGATTTCTAACAGGAGAGC 951
Qy 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
Db 952 CATAGACATTTCCAAAGAGTATATTGTAGACGACGAGATGTGGAGACGAGATGAATAT 1011
Qy 361 GluArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
Db 1012 GAAAGCAAGAGAGAGAGGAGGTACCAAGGACGCTACAGAGTGTATCCAAATTTGGCC 1071
Qy 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
Db 1072 CGGTATCCGGTAAAGCCACAACTTATGAAGCAAAATGGGATCCACGCTGAAGTGTC 1131

Qy 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
Db 1132 CGGCACACACAGAGAGAGGACAGTATGTTCTTTGGCAATGCTGAACCTAGAAGAT 1191
Qy 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
Db 1192 TCCAGATTCTCTGCTAAGGATGATAGACCATCAGGCAAGATCTGATCTGAACGT 1251
Qy 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
Db 1252 AGAGCTGCAATGGAAAACACAGCATCGTATTCAATGGAAAGAACTCGAAGAGCTCAGGA 1311
Qy 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
Db 1312 CAAGTTCCTTATCCAAAGGACCAAAATCATAGTCTCTTACCCCTCGCAGAGGCCCT 1371
Qy 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLysLeu 500
Db 1372 ATACCGCTCGATAGACAGAGCTGAGCGTGGCGACTCCCTACGGAACACACACCTTA 1431
Qy 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
Db 1432 GATCCAGCTCTGCTGTAAGGAAAACGAAAGCGAGAAAATAATGAAACAAATGTTGAGGAAT 1491
Qy 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProArgProHisLys 540
Db 1492 GATTCCTTGAGTTCAGACAGTCTGAGTCAGTAGGCGCCGCCACCAAGGCCCTCATAAA 1551
Qy 541 SerLysLysGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAla 560
Db 1552 TCCAAAGAGAGGTTAAATGGCCAGCTTTCACCTGACGAGCTCAGAGGAGGTTGGCA 1611
Qy 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
Db 1612 TCCACGCTCGATATACAAAGCTGTGATGACGTGAGATGAAAGCGAGAGGTAGTGAG 1671
Qy 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1672 AAAGAGACATGGAG----- 1686
Qy 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1687 -----TACAGCTGGTTGGAG 1701
Qy 621 AspLeuGluTyrSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 640
Db 1702 CATGCTCTTGGCAT----- 1716
Qy 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1717 -----AGCAGTAGGCATCCCAATGCTTTGCAACCTGTGACCTGGCAGCATCCAAA 1770
Qy 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 1771 GATGAGATCGCTAAATTTGGTTCGTATTATTAATAAGCGTTTAAAGATGGAGTGTA 1830
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLysValValGlyLysMetThrGluSer 700
Db 1831 CCTCAGATTCAGGAGCAATCTGGGCTTAAAGGTTGTAGGAGGAAAGATGACTCAATCA 1890
Qy 701 GlyArgLeuCysAlaPheIleThrLysValLysGlySerLeuAlaAspThrValGly 720
Db 1891 GGTCCACTTTGTGCATTTATTACCAAGTGAAGAAAGAGAGTTAGTGTACTGTAGGA 1950
Qy 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 1951 CATCTTAGACAGGTCATGAAGTCTTGGAAATGGAAATGGAGGCTATTCCAGGAGCCACA 2010
Qy 741 PheGluGluValTyrAsnIleLeuGluSerLysProGluProGlnValGluLeuVal 760
Db 2011 TTTGAGGAAGTTTACCAATATTCTAGATCCCAAGCTGAAACCAACCAAGTTGAGCTTGT 2070

761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
2071 GTTTCAGGCGGATGAGATATGCTAGATACCTGATAGCACTCATGACACACTGGA 2130
781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
2131 TCCAGATTCTAGCTCATTTGAATCTCAAAAAATGATCGTCCGCTATATCGGTACCTCT 2190
801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
2191 CCATGAGTCTCGCATGCTGAGGAGATGCCACAGTTCATTCGACACGCTTCAATA 2250
821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
2251 AACTATGTTTGAAGGTTGGTCCACAGTTATAGTTACATTTTGGAGGCAAGGAT 2310
841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
2311 CTCCCTTCCAGGAAGATGGAGGCCAAGGAATCTTTATGTTAAATTTACTTCTTCCA 2370
861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
2371 GACAGAGTATATAAACAAGAGAGAACAAAAACAGTCAAGAAAACTTTGGAAACCCAAA 2430
881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
2431 TGGNACCAAGCTTTCATTTATTTCTCCGTTCCACGAAGAGAAATCCGGAGCAAGATGCTC 2490
901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920
2491 GAAATCACTTTGGGATCAAGCTCGAGTTCGAGAGGAAGAAAGTGAATTCCTAGGAGAG 2550
921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGln 940
2551 ATTATTAATGAAATGGAACAGCTCTGTTAGATGAACACACACTGGTACAACTTCAG 2610
941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
2611 ACCATGATGTCCTCATTTGCCACTCCCTCACCTCTCCATATATGCCAGGAGACAG 2670
961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
2671 CTGCATGAGAGAGCCCAACACGAGGTTGCAAGGTCGAAGAGAAATAGTGACAGCGAA 2730
981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
2731 GTATCCGACTATGACTCGAGGAGCGCGTGGAGTCTGTGATTAACGACACGATGGC 2790
1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
2791 CGTGATCTTCAAGCTCCATATATCATGTCGACGACAAAGTCATGTCATCAAAACCATTCG 2850
1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
2851 TCACCATCAGCTCTCTCATCGATAGATGTTATAGGAAGGACTAGTCTATGTCGCTCT 2910
1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
2911 AGTGTCCTCTCTCAAGGAATGTGAAACAGGGGCTTCGAGGACACGCTGCTACTGCGC 2970
1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
2971 CATTAACATAATAGCCGATGGACAGACACCGTGTCTATGGACGACCATTACTTCTCA 3030
1081 AspArgAsp----- 1083
3031 GAGAGACAGCTCATTTTCTCACTCTACTCTGCTCTCGACACAGGACAGCAGTGACAT 3090
1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
3091 CACCACAGGGATGGAAGGATTTGTAAGCAGCAGATAGACAGCCATATCATCAGATCCAGA 3150
1099 SerThrGluGlnArgProLeuLeuGluArgThrThrArgSerArgSerGluArg 1118

3151 TCACAGAGACACAGCGCTCTCTTAGAGCGGACCAACCCGCTCCAGATCTCTTGACGT 3210
1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
3211 GCTGATACAAACCTCATGAGGTGATGCTTCAATTAATGACTGGAAGATCTGCCCTCCT 3270
1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
3271 TCACCTGCTCTATCGAGGTCTCACCTCGACCTGGCTCTGTCCAGACCACTGCGTCAAGT 3330
1159 ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProLysGlyThrLeu 1178
3331 ACTCCGTAACAGACAGAGGGCCCGCAGCTTCTCAGCTCCACCAAGGGAACACTG 3390
1179 GluArgSerAlaMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrLys 1198
3391 GAAAGA----- 3396
1199 GlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGly 1218
3396 ----- 3396
1219 TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSer 1238
3396 ----- 3396
1239 AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMet 1258
3396 ----- 3396
1259 SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet 1278
3396 ----- 3396
1279 GlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGly 1298
3396 ----- 3396
1299 AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu 1318
3396 ----- 3396
1319 GlyThrSerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGly 1338
3396 ----- 3396
1339 LeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLys 1358
3397 -----GAGAGGTTAAAG 3411
1359 LeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrp 1378
3412 CTACGACGACCTGCTCCAGAGACAGCAGAGCCGGCTAGCAGTGCAGATGAGAACTGG 3471
1379 MetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGlu 1398
3472 ATGACAAGACAGCCACCGCGAATCCACGATGCGACGATGACAGCTACAGCTCAGAA 3531
1399 GlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeu 1418
3532 GGAATCTGATCTTCTCGGGTCCGCTCTGACAGCCAGTTCAGCGATTCCTT 3591
1419 AspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly 1438
3592 GATGTCCTGGCCCTGCTCAGCTAGTGGAGCGCAGACCTCGGACTCTCTGCAATGGGT 3651
1439 AspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleLeuArgAla 1458
3652 GACATTCAGTGGGATGATGATAAGAGGAGCAGCTGAGGTAGAGATCATCCGGCT 3711
1459 ArgGlyLeuValLysProGlySerLysThrLeuProAlaProTyrValLysValTyr 1478

Db 3712 CGAGGCTTGTAGTAAACAGGCTCCAGACACTGCCAGCACCATATGTCAGGTGTAC 3771
Qy 1479 LeuLeuAseAseGlyValCysIleAlaLysLysThrLysValAlaAseGlyThrLeu 1498
Db 3772 CTGTGTAGACACAGGAGTGTGCATAGCAAAAGAAACCAAGGTGGCGAGAAAGCGCTG 3831
Qy 1499 GluProLeuTyGlnGlnLeuLeuSerPheGluGluSerProGlnGlnArgValLeuGln 1518
Db 3832 GAGCCCTGTATCAGCAGCTGTATCTTTCGAGAGAGAGCCCAAGGAGGTGTACAG 3891
Qy 1519 IleIleValTTPGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGln 1538
Db 3892 ATCAATGTCTGGGAGATTATGCGCGCATGGATCAAAATCCTTATGGAGTGGCCAG 3951
Qy 1539 IleLeuLeuAspGluLeuLeuSerAsnMetValIleGlyTTPheLysLeuPhePro 1558
Db 3952 ATACTCTTAGAGTGAATGGAATATCAACATGGTAATGGATGTTCAAACTCTTCCT 4011
Qy 1559 ProSerSerLeuValAspProThrSerAlaProLeuThrArgAlaSerGlnSerSer 1578
Db 4012 CTTTCTCTCTAGTAGATCAACCTTGGACCTCTGACAGAGAGCTTCCCAATGCT 4071
Qy 1579 LeuGluSerSerThrGlyProSerTyrSerArgSer 1590
Db 4072 CTGGAAGTCTACTGGACCTTCTACTCTCGTTCA 4107

RESULT 10

AF199326
LOCUS AF199326 5031 bp mRNA linear ROD 04-JUL-2000
DEFINITION Rattus norvegicus RIM2-3B (Rim2-3B) mRNA, complete cds.
ACCESSION AF199326
VERSION AF199326.1 GI:8925865
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 5031)
AUTHORS Wang, Y., Sugita, S. and Sudhof, T.C.
TITLE The RIM/NIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins
JOURNAL J. Biol. Chem. 275 (26), 20033-20044 (2000)
MEDLINE 20347919
PUBMED 10748113
REFERENCE 2 (bases 1 to 5031)
AUTHORS Wang, Y. and Sudhof, T.C.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-1999) Center for Basic Neuroscience, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, TX 75235-9111, USA

FEATURES

source
1..5031
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
gene
1..5031
/gene="Rim2-3B"
CDS
52..4110
/gene="Rim2-3B"
/codon_start=1
/product="RIM2-3B"
/protein_id="AA081648.1"
/db_xref="GI:8925866"
/translation="MSAPLGRGRPAFTPAASQPPPPQPMPLSHLTBEERKIIQAVM
DRQKEEKEQSVLKKLHQFMYKQVKWGEESQOQEQKGDAPTGGCHTKTFAD
GGHNCVCTQKFCACGGRVSLRKNWVNCNCRKQELITKSGAPYNGSFTPO
QDQKALGRSEAPQEKAKLHEQTQGFPGDSVPAVERGRAGHTRQSLNG
SGKKHQAIDMPSDRKRSVSDQNRRYDQSEEREYSQYPSDTPRSPSDYADR
RSQRFQFYFVYHLNDRNRGRHSHKEYIVDDVEDSRDEYERQREERYQARYR
SDPNLARVPKQPYEQMRHAEVSRARHRRHSDVSLANAELEDSRISLMDRPS
RORSVSEREAAMENORSYMERTRAQGSSYPQRTNHSPTPRSPPTPLDRPELR
ADSLRKQHLDPSSAVRKTREKMETMLRNDLSLSDQSESVPVPPPPPHKSKKGGMR

QVLSSEELASTPEVTSDDVEIESGVEKGMVEYSWLEHSMHSSSEASPMSLHP
VTWPSKXGDELIGRIILNKLKDGSPROSGMGLKXVCGEMTESGRICAFITKXK
KGSADTVGHLRPGDEVLEWNGRLLOCATFEVYNIILSRPEQVELVSRPQDMF
RIPDTTQALLESSSSSFESQMDRPSISVTPSPGMLRDVPQFSLGSLIKWPFDKV
GHQIVITLGAADLPSEDEGRPNFYKIFLPDRSDKRTKTKTKLTPKXNQT
IYSPHLRPFERMELEITLWDQARVEEESFLEIILELTALLDDEPHWYKQTHD
VSSPLPHSPMPRROLHGSPTRRLQSRKISDSVSDYDCDGVGVVDYRHDGR
DLQSTLSVPEQVMSNHCSPSGSPHVDVIGTRTSPSPSPVPPQRNVQGLCTRAT
GHYNTISRMDEHRVMDHYSSEROSHFTLPRSRHRTQSEHHHRDGRDCEAARQPHY
KSRBQRPLLETRITTSRERADTNLMRSPSLMTSRAPPSPALSRPTGTSVQ
TSPSTPTVTRGRQLPQPKGLRGGKRLSTVQSTETGLAVEMRNWMTQAS
RESTDGSNYSSEGNLIFPQRLASQSFDFLDGLGPAQLVGRQTLATPMDGIOV
GMWDKGLQLEVEIIRAGLVVPGSKTLPAPYKVLVLLDNGVCTAKKTKVARTLEP
LYQQLLSEESPOQKVLIVKPGDYGRMDHKSFMGVAQLLDELLELSNWTGWKLPFP
PSSLVDPLAPLTPRASQSSLSSTGTSYRS"

BASE COUNT 1503 a 1166 c 1246 g 1116 t
ORIGIN

Alignment Scores:

Pred. No.: 7,72e-197 Length: 5031
Score: 6522.00 Matches: 1283
Percent Similarity: 80.71% Conservative: 18
Best Local Similarity: 79.59% Mismatches: 29
Query Match: 78.72% Indels: 282
DB: 10 Gaps: 5

US-09-617-099B-1 (1-1590) x AF199326 (1-5031)

Qy 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAseGlnPro 20
Db 52 ATGTGGCTCCACTCGGGCCCGGGCGCGCGCTCCACCCCGGGCGCTCTCAGCCT 111
Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIle 40
Db 112 CCGCGAGCCGAGATCCGAGCTCAGCACCTCAGCGAAGAGAGGAGGAATCATC 171
Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysIle 60
Db 172 CAGGCTGTATGATCGTCAGAGAAAGAGAGAGGAGGAGCGATCGTCTCAAA --- 228
Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 228 ----- 228
Qy 81 ValAsnAsnValLeuGlnProGlnLysGlnProAsnGluLysGluProGlnThrLys 100
Db 229 -----AAG 231
Qy 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
Db 232 CTGCATCAACATTTGAATGTATAGGAGCAAGTCAAGAAATCGGAGAGGAATCACA 291
Qy 121 GlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
Db 292 CAGCAGCAAGAGCAGAGGCGCGCGCGCGCATCTGCGCAACAGCAAAATTT 351
Qy 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
Db 352 GCAGATGGATCGCGCCATAACTGTTTCAATATGCCAACCAAGTCTGTGCTGTGTGGA 411
Qy 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
Db 412 GGTGAGGTGTATATACGCTCAACAGGTATGTGGGTGTGTATTTGTGCGCAAAACA 471
Qy 181 GlnGluIleLeuThrLysSerGlyValAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
Db 472 CAAGAAATCTCACTAAGTCCGGCGCGGTGTTTACATAGTGGGTCTAATACCCGAG 531
Qy 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluAlaProGlnGlnLysLys 220
Db 532 CAGCCTGATCAAAAGGCTCTTCGAGGGCTTCGAGTGAAGAGGCCCTCAGGAGAGAAG 591
Qy 221 AlaLysLeuHisGluGlnProGlnPheGlnAlaProGlyAspLeuSerValProAla 240

592 GCACAACTGATGAGCAGACGAGTCCAGGACCCCGGTGACTCATCAGTACCTGCA 651
241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
652 GTTGAGAGAGGTGAGCTCATGGCTGACACAGACAGGATTCATTAAAGATGGCTCAGGA 711
261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
712 ATGAGCACCAGATCCGAGTGCATGCTTCACAGCAAAAGAAAGATCGTCAGTGTCC 771
281 ArgAspGlnAsnArgArgGlyGluGlnSerGluGluArgLysArgSerGlnIleVal 300
772 AGGGATCAAAATCGAAGATACGACCAAGAGTGAAGAAAGAGAGGAATATTCCAGTATGT 831
301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
832 CCTTCAGATAGCAATGCTGATCTCCATCAGATTATGCTGATAGCATCTCAACGT 891
321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
892 GAACTCAATTTATGAGACCTGATCATTTAAATATTAGGATTCACAGAGAGGC 951
341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
952 CATAGACATTCACAAAGATATATTGTAGACGACGAAGATGTGAGAGCAGAGATGAAT 1011
361 GluArgGlnArgArgGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
1012 GAAAGCAAGAGAGAGAGGAGTACCGACGCTACAGAAAGTATCCAAATTTGGCC 1071
381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
1072 CGGTATCCGGTAAAGCCACACCTATGAGACCAATGCGGATCCACGCTAAGTGTCC 1131
401 ArgAlaArgHisGluArgHisSerAspValSerLeuAlaAsnAlaGluLeuAsp 420
1132 CGGGCACGACACGAGAGAGGACAGTGTATTTTGGCAATGCTGAATAGAGAT 1191
421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
1192 TCAGGATTTCTGTCTAAGGATGGATAGACCATCAAGGCAAGATCTGTATCTGAACGT 1251
441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
1252 AGAGCTGCATGAAACAGCGATCGTATTCATGGAAGAACTCGAGAAGCTCAGGA 1311
461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
1312 CAAAGTCTTATCCACAAAGGACCAAAATCATAGTCTCTTACCTCGCAGGAGCCCT 1371
481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
1372 ATACCGCTCGATAGACGAGTGGCGGTGCGACTCCCTACGGAAACACACCACTTA 1431
501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
1432 GATCCAGCTCTGTGTAGGAAACAGAGCGGAGAAAATGGAACAAATGTGTAGGAT 1491
521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
1492 GATTTCTTGAGTTTCAGACGATCTGAGTTCAGTGGCCGCCCCACCAAGGCTCATAAA 1551
541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560
1552 TCCAGAAAGAGGAGTAAATGCGCCAGGTTTCACTGAGCAGCTCAGAGGAGGATGGCA 1611
561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
1612 TCCAGCCCTGATATACAGCTGTGATGACGTGGAGATTTGAAAGCAGAGCGGTAGGTAG 1671
581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
:::
1672 AAAGGAGACATGAG-----1686
601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyLysHisSerLeuGluGlu 620
1687 -----TACAGCTGGTTGGAG 1701
621 AspLeuGluTyrSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
1702 CATGGCTCTGGCAT-----1716
641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
1717 -----AGCAGTGAGCATCCCAATGTCTTCCACCTGTGACCTGGAGCGCATCCAAA 1770
661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLysAspGlySerVal 680
1771 GATGAGATCGCTTAATTGGTGTATTTTAAATAAGCTTAAAGATGGAGTGA 1830
681 ProArgAspSerGlyValaMetLeuGlyLysValValGlyGlyLysMetThrGluSer 700
1831 CTTGAGATTCAGGAGCAATGCTGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA 1890
701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
1891 GGTGACCTTTGTCATTTATTACCAAGTGAAGAAAGAGAGTTTACTGATCTAGTGA 1950
721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuGluGlnGlyAlaThr 740
1951 CATCTTAGACCGATGATGAGTCTTGAATCGAATGGAAGCTATTGCAAGGAGCCACA 2010
741 PheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVal 760
2011 TTTGAGAGATTTACCAATTTCTAGAATCAAGCTGAAACCAACCAAGTTGAGCTGT 2070
761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
2071 GTTTCAGGCGGATTTGGAGATATGCTAGATACCTGATAGACATCTGTCACACTGGAA 2130
781 SerSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
2131 TCCAGTCTTAGCTCATTTGAAATCTCAAAAATGGATGCTGCTCTATATCGGTACCTCT 2190
801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
2191 CCCATGAGTCTCGCATCTGAGGAGATCTCCACAGTCTTATCTGCAAGCTTCAATA 2250
821 LysLeuTyrPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyValLysAsp 840
2251 AAACATATGTTTGACAAGTTGGTCCAGATTATAGTTACAAATTTGGGCAAGGAT 2310
841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
2311 CTCCCTTCAGGAGATGGGAGGCAAGGAATCTTATGTTAAATTTACTTCTTCCA 2370
861 AspArgSerAspLysAsnLysArgThrLysThrValLysLysLysLeuIleuLysProLys 880
2371 GACAGAGTGTAAACACAGAGAGAAACAAACAGTCAAGAAACCTTTGGAAACCCAAA 2430
881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
2431 TGGAAACAGACTTTCATTATTCTCTCTGTTTCCAGAGAGAAATCCGGGAGCAATGCTC 2490
901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGlnPheLeuGlyGlu 920
2491 GAAATCACCTTTGGGATCAAGCTGAGTTCGAGAGGAGAAAGTGAATTTCTTAGGAGAG 2550
921 IleLeuIleGluLeuGluThrAlaIleLeuAspAspGluProHisTyrLysLeuGln 940
2551 ATTTTAATTTGAATTTGGAACAGCTCTGTAGATGATGAACACACACTGGTCAAACTTCAG 2610
941 ThrHisAspValSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
2611 ACCCATGATGCTCTCTTATGCCACTCCCTCACCTTCTCCATATATGCCACGGAGACAG 2670

QY 961 LeuHisGlyGluSerProThrArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
 DB 2671 CTGCATGGAGAGAGCCCAACGCGAGGTTGCAAGGTCGAGAGATTAAGTGCAGCCAA 2730
 QY 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
 DB 2731 GTATCCGACTATGACTCGAGGACCGCGTGGAGTCGTGCAGATTACCGACACGATGC 2790
 QY 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
 DB 2791 CGTGATCTTCAAGCTCCATATATAGTCCAGAACAAATGTCATCAAAACCATTC 2850
 QY 1021 SerProSerGlySerProHisArgValAspValIleGlyValArgThrArgSerPro 1040
 DB 2851 TCACCATCAGCTCTCTCATCGAGTAGATGTTATAGAAAGACTAGGTCAATGTCGCT 2910
 QY 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
 DB 2911 AGTGTCCTCTCTCTCAAGGAATGTGAACAGGGGCTTCGAGGACACGTGCTACTGCG 2970
 QY 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
 DB 2971 CATTAACAATCAATTAGCCGATGGACAGACACCGTGTCTATGGACCACTTACTCTTCA 3030
 QY 1081 AspArgAsp----- 1083
 DB 3031 GAGAGAGACAGTCATTTTCTACCTCTACCTCGCTCTCGACAGCGACCACTGAGCAT 3090
 QY 1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
 DB 3091 CACCACAGGAGTGAAGGGATTGTGAAGCAGCAGATAGACAGCATATCACAGATCCAGA 3150
 QY 1099 SerThrGluGlnArgProLeuLeuArgThrThrThrArgSerArgSerGluArg 1118
 DB 3151 TCAACAGAACAAACGGGCTCTCTTAGAGCGGACCAACCCGCTCCAGATCTCTGAACGT 3210
 QY 1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
 DB 3211 GCTGATACAACTCATGAGTCGATCGCTTCTTAATGACTGGAAGATCTGCCCTCTCT 3270
 QY 1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
 DB 3271 TCACCTGCCTTATCGAGTCTCACCTCGCTCTCGCTCTCTCGACAGCATCTCGTCAAGT 3330
 QY 1159 ThrProGlyThrGlyArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeu 1178
 DB 3331 ACTCCCGTAACAGGACGAGGGGCGGAGCTTCTCTCAGCTCCCAACAAAGGGAACACTG 3390
 QY 1179 GluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLys 1198
 DB 3391 GAAAGA----- 3396
 QY 1199 GlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGly 1218
 DB 3396 ----- 3396
 QY 1219 TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerValSer 1238
 DB 3396 ----- 3396
 QY 1239 AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMet 1258
 DB 3396 ----- 3396
 QY 1259 SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet 1278
 DB 3396 ----- 3396
 QY 1279 GlnAsnArgGlnMetGlyValSerGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGly 1298
 DB 3396 ----- 3396

QY 1299 AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu 1318
 DB 3396 ----- 3396
 QY 1319 GlyThrSerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGly 1338
 DB 3396 ----- 3396
 QY 1339 LeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLys 1358
 DB 3397 -----CGAGGAGTAAAG 3411
 QY 1359 LeuArgSerThrValGlnArgSerThrThrGlyLeuAlaValGluMetArgAsnTrp 1378
 DB 3412 CTACGCGACACACTCTCCAGAGAGACGAGGACCGGCTAGCAGTGGAGATGAGGAACCTGG 3471
 QY 1379 MetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGlu 1398
 DB 3472 ATGACAAGACAGCCAGCCGCGGAATCCACGATGGCAGCATGAACAGTTACAGCTCAGAA 3531
 QY 1399 GlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeu 1418
 DB 3532 CGAAATCTCATCTTCTCGGGTCCGCTCTGACAGCCAGTTCAAGCATTTCTCTT 3591
 QY 1419 AspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly 1438
 DB 3592 GATGGTCTGGGCTCTGCTCAGCTAGTGGACGCCAGACCTCGGCACTCTCAATGGGT 3651
 QY 1439 AspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAla 1458
 DB 3652 GACATTCAGTGGGATGATGATTAAGAGGACAGCTGGAGGTAGAGATCATCCGGGT 3711
 QY 1459 ArgGlyLeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyr 1478
 DB 3712 CGAGGCTTGTAGTAAACCAAGTCTCAAGACACTGCCAGCACCATATGTCAAGGTGTAC 3771
 QY 1479 LeuLeuAspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeu 1498
 DB 3772 CTGTAGACACGGAGTCTGCATAGCAAAAGAAACCAAGGTGGCGAAGACGCTG 3831
 QY 1499 GluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGln 1518
 DB 3832 GAGCCCTGTATCAGCAGCTGTATCTTTCGAGAGAGAGCCGCCAGGGAGGTGTACAG 3891
 QY 1519 IleIleValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGln 1538
 DB 3892 ATCAATGTCTGGGAGATTAATGCGCATGGATCAAAATCTTTATGGAGTGGCCAG 3951
 QY 1539 IleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPhePro 1558
 DB 3952 ATACTCTTAGATGAATCGAATATCCCAATGGAATGGATGGTTCAAACTCTTCCCT 4011
 QY 1559 ProSerSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSer 1578
 DB 4012 CTTTCTCTCTAGTAGATCCAACTTGGCACTCTGACAGAGAGAGTTCCTCCATCGTCT 4071
 QY 1579 LeuGluSerSerThrGlyProSerTyrSerArgSer 1590
 DB 4072 CTGGAAGTCTTACTGGACCTTCTTACTCTCTCTCA 4107

RESULT 11

LOCUS AF199327
 DEFINITION Rattus norvegicus RIM2-4A (Rim2-4A) mRNA, linear ROD 04-JUL-2000
 ACCESSION AF199327
 VERSION AF199327.1 GI:8925867
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 5031)

QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
DB 1192 TCAGGATTTCTCTGCTAGGATGGATAGACCATCAAGGCAAGATCTGTATCTGAACGT 1251
QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
DB 1252 AGAGCTGCAATGGAACACAGCGATCGTATTCATGGAAGAACTCGAGAAGCTCAGGGA 1311
QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
DB 1312 CAAAGTTCCTATCCACAAAGACACCAAAATCATAGTCTCTCCCTCGCAGAGCCCT 1371
QY 481 IleProLeuAspArgProAspMetArgAlaAspSerLeuArgGlnHisLeu 500
DB 1372 ATACCGCTCATAGACAGAGCTGAGCGCTGCCGACTCCCTACGGAACACACCACTTA 1431
QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
DB 1432 GATCCAGCTCTCTGTAGGAAACGACGAGGAGAAATGGAACATGTTGAGGAAT 1491
QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
DB 1492 GATTCCTTGTAGTTACAGCAGCTCTGAGTCACTGAGCGCGCCGCCACCAAGGCCCTCATAA 1551
QY 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560
DB 1552 TCACAGAAAGAGGTAAATGCCAGGTTTCTAGCGAGCTCAGAGGAGGAGTTGGCA 1611
QY 561 SerThrProGluTyrThrSerCysAspValGluLeuGluSerGluSerValSerGlu 580
DB 1612 TCACGCTGAGTATACAGCTGATGAGTGGAGATGAAGCGAGAGCGTAGTGAG 1671
QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
DB 1672 AAAGGACATGCG----- 1686
QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGlu 620
DB 1687 -----TACAGCTGGTGGAG 1701
QY 621 AspLeuGluTrpSerGluProGlnIleLysAspSerGlyValAspThrCysSerThr 640
DB 1702 CATGCGTCTGGCAT----- 1716
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
DB 1717 -----AGCAGTGGAGCATCCCAATCTCTTTCACCCCTGTGACCTGGCAGCCATCCAA 1770
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
DB 1771 GATGGAGATCGCTTAATTTGCTGTATTTTATTAATAAGCGTTTAAAGATGGAGTGTA 1830
QY 681 ProArgAspSerGlyAlaMetLeuGlyLysValValGlyGlyLysMetThrGluSer 700
DB 1831 CCTCGAGATTCAGAGCATGCTGGCTTAAAGTTGTAGGAGAAAGATGACTGATCA 1890
QY 701 GlyArgLeuCysAlaPheIleThrLysValLysGlySerLeuAlaAspThrValGly 720
DB 1891 GGTGCACTTGTGCAATTTATTCACAAAGTGAAAGAAAGTGTAGCTGATCTGTAGGA 1950
QY 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
DB 1951 CATCTTAGACCGGTGATGAAGTCTTGGATGATGAATGGAAGGCTATTGCAAGGAGCCACA 2010
QY 741 PheGluGluValTyrAsnIleLeuGluSerLysProGluProGlnValGluLeuVal 760
DB 2011 TTTGAGGAGTTTACACATTTCTAGAAATCCAGCTCAAGCTCAACCAAGTTGAGCTTGT 2070
QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
DB 2071 GTTTCAGGCGGATGGAGATATGCTTAGAATACCTGATGACACTCATGCACAACTGGAA 2130
QY 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800

DB 2131 TCAGTTCAGCTCATTTGAAATCTCAAAANATGATGTCCTCTATATCGGTACCTCT 2190
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
DB 2191 CCCATGAGTCTGGCATGCTGAGGAGATGTCACAGTTCCTATCTGGACAGCTTCAATA 2250
QY 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
DB 2251 AAACATATGTTTGGACAGGTTGGTCACCAAGTTAATAGTTTACAAATTTGGGACCAAGAT 2310
QY 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
DB 2311 CTCCCTTCAGGAGATGGGAGGCCAGGAATCTTATGTAAATTTACTTCTTCCA 2370
QY 861 AspArgSerAspLysAsnLysArgThrLysThrValLysLysThrLeuGluProLys 880
DB 2371 GACAGAGGTGATAAAACAAAGAGAGAACAAACACAGTCAAGAAACCTTTGGACCCAAA 2430
QY 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
DB 2431 TCGAACCCAGACTTTCATTTATTTCTCTGTTTCCACGAGAGAAATTCGGGAGCAATGCTC 2490
QY 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920
DB 2491 GAATTCACCTTTGGATCAGCTCGAGTTCCAGAGGAGAAAGTGAATTTCTAGAGAG 2550
QY 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGln 940
DB 2551 ATTTTAATTTGAATGAAACAGCTCTGTTAGATGATGAACACACTGTTACAACTTCAG 2610
QY 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
DB 2611 ACCCATGATGTCCTCATTTGCTCCCTCCCTTCCATATATATATATATATATATATAT 2670
QY 961 LeuHisGlyLysSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
DB 2671 CTGCATGGAGAGAGCCCAACACGAGGTGCAAGGTCGAAGAGATAAGTACAGCGGAA 2730
QY 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
DB 2731 GTATCGCATATGACTCGCGAGCGCGCTGGGAGTCTGTCTAGATTCACGACGATGCG 2790
QY 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
DB 2791 CGTGATCTTCAAAAGCTCCACATTCAGTCCAGACCAAGTCTATGTCTATCAAAACCATTC 2850
QY 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
DB 2851 TCACCATCAGGCTCTCTCATCGAGTAGATGTTATAGGAAGGACTAGGTGCTGCTGCT 2910
QY 1041 SerAlaProProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
DB 2911 AGTGTCTCTCTCTCAAGGAATGTGGAACAGGGGCTCGAGGAGACAGTGTACTGCG 2970
QY 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
DB 2971 CATTAACAATACCAATAGCCGATGGACAGACACCGGTCTATGACGACCATTAATCTTCTCA 3030
QY 1081 AspArgAsp----- 1083
DB 3031 GAGAGAGACAGTCATTTTCTACTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3090
QY 1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
DB 3091 CACCACAGGATGGAGAGGTGTGAGCAGCAGATAGACACCCATATTCACAGATCCAGA 3150
QY 1099 SerThrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerGluArg 1118
DB 3151 TCAACAGAAACACGGCTCTCTCTAGAGCGGACCAACCCCGCTCCAGATCTCTCTGAACGT 3210
QY 1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138

IYSPVRRERFRMLEITLWDOARVRESESEPLGBILTELETALELDDPHWKLQTHD
 VSSLPLPSPWPRQHLGESPTRRLORSKEISDSSEVSDYDCBGVGVSDYRDRG
 DLQSTLSVPEQVMSNHCSPSGSHRVDVIGRTRSNPSFPVPPORNYEQGLRGTAT
 GHYNTISMDRHRVMDHYSSERDSHFLTPRSRHRQTSERHHRDRDCZEAADROPTH
 KRSRTEQRLPRLTRRSERADTNLMRSMPLTGRSAPPSPALSRSHPTGVSQ
 TSPSTPTVGRGROLPLPPKGLTGERGGKLRSTVORSTETGLAVRNWMTQAS
 RESTDGSNMSYSEKMLIPPGVRLASDQFSDPLDGLPAQLVGRQTLATPAMGDIQV
 GMDXKQGLEVILIRARGLVKGSKTLPAPYKVYLLDNGVCIACKTKVARKTLEP
 LFQQLSEESFQKRVLIWVDYGRMDHKSFMGVAQILLDELSNWNVGNFKLFP
 PSSLVDPTLAPTRASQSLSSSTGSPYSRS*

BASE COUNT 1503 a 1166 c 1246 g 1116 t

ORIGIN

Alignment Scores:

Pred. No.: 7,72e-197 Length: 5031
 Score: 6522.00 Matches: 1283
 Percent Similarity: 80.71% Conservative: 18
 Best Local Similarity: 79.59% Mismatches: 29
 Query Match: 78.72% Indels: 282
 DB: 10 Gaps: 5

US-09-617-099B-1 (1-1590) x AF199332 (1-5031)

QY 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 DB 52 ATGTCGGCTCCACTCGGCGCCCGGGCCGCGCGCTCCACCCCGGGGCTCTCAGCCT 111
 QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIlelle 40
 DB 112 CCTCGCAGCCGAGATCGGACCTCAGCCACCTCAGGAAAGAGAGAGAGAAATCATC 171
 QY 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysIle 60
 DB 172 CAGGCTGTCTGATGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
 QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
 DB 228 ----- 228
 QY 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
 DB 229 -----AAG 231
 QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
 DB 232 CTGCATCAACAATTTGAAATGTATAAGGAGCAAGTCAAGAAATGGGAGAGGAATCACAA 291
 QY 121 GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
 DB 292 CAGCAGCAAG 351
 QY 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
 DB 352 GCAGATGGATGCGGCCCAACTTTCATATTCGCAACCAAGAGTCTGTGTCTGTGTGGA 411
 QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
 DB 412 GGTGAGTGTATCATCGCTCAACAAGAGTATGTGGGTGTGTATTTGTGCGCAAAACAA 471
 QY 181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
 DB 472 CAGAATCTCTACTAAGTCGGCGCGGTGTGTTTACATAGTGGTCTAATACACCGCAG 531
 QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluAlaProGlnGluLysLys 220
 DB 532 CAGCCTGATCAAAAGGCTCTTCGAGGCTTCGAGTGGAGAGAGAGAGAGAGAGAGAG 591
 QY 221 AlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
 DB 592 GCAAAACTGATGAGCAGAGCGAGTTCAGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
 QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260

DB 652 GTTCAGAGAGGCTCGAGCTCATGGGCTCACAAGACAGGATCTTATTAAAGATGGCTCAGGA 711
 QY 261 VallyHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 DB 712 ATGAAGCACCAGATCGCCAGTGACATGCTTTACAGACAGAAAAGAGTCCCTCAGTGTC 771
 QY 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
 DB 772 AGGATCAAAATCGAAGATACACCAAGTGAAGAGAGAGAGAGATATTTCAGATATGTT 831
 QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
 DB 832 CCTTCAGATAGCACAAATGCTTAGATCTCCATCAGATTTATGCTGATAGACGATCTCAACGT 891
 QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
 DB 892 GAACCTCAATTTTATGAAGAACCTGATCATTTAAATTTATAGGGATTTCTAAACAGGAGAG 951
 QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
 DB 952 CATAGACATTTCCAAAGAGTATATTGTAGACGACGAGATGTGGAGAGCAGAGATGAAT 1011
 QY 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
 DB 1012 GAAAGACAAAGAGAGAGAGAGAGTACCAGGACGCTACAGAGTGTATCCAAATTTGGCC 1071
 QY 381 ArgTyrProValLysProGlnProTyrGluGlnGlnMetArgIleHisAlaGluValSer 400
 DB 1072 CGGTATCCGGTAAAGCCCAACCCCTATGAAGAGCAAAATGCGGATCCACGCTGAAGTGC 1131
 QY 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuAsp 420
 DB 1132 CGGGCACACAGAGAGAGAGAGCAGTGTATTTCTTTGGCAATGCTGAATAGAGAT 1191
 QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
 DB 1192 TCCAGATTTCTCTGCTAAGGATGATAGACCATCAGGCAAGATCTGTATCTCAACGT 1251
 QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
 DB 1252 AGAGCTGCAATGGAAGAACCCAGCGATCGTATTTCATGGAAGAACTCGAGAGAGCTCAGGA 1311
 QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
 DB 1312 CAAAGTTCTTATCCAAAGGACCAAAATCATAGTCTCTTACCCCTTCGCGAGAGCCCT 1371
 QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
 DB 1372 ATACCGCTCGATAGACAGAGCTGAGGCGTGCAGCTCCCTACGGAACCAACACCACTTA 1431
 QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgGln 520
 DB 1432 GATCCAGCTCTGCTGTAAAGGAAAACGAGCGAGAAAATAATGGAACCAATTTGTAGGAAT 1491
 QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
 DB 1492 GATTCTTTGAGTTTCAAGCCAGTCTGAGTCAGTGAGCGCCGCCCCACAGGCTCATAAA 1551
 QY 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560
 DB 1552 TCCAGAAAGAGAGGTAATAATGCGCAGGTTTCACTGAGCAGCTCAGAGAGAGTGGCA 1611
 QY 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
 DB 1612 TCCAGCCTCGATATACAGCTGTATCAGCTGAGATTCAGAGCGAGAGAGAGAGAG 1671
 QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValIleuSerAspSer 600
 DB 1672 AAAGAGACATGGAG----- 1686
 QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGlu 620
 DB 1687 -----TACAGCTGGTGGAG 1701

QY 621 AspLeuGluTrpSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
Db 1702 CATGGCTTGGCAT----- 1716
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1717 -----ACGAGTGAGGCATCCCAATGCTTTTGACACCCCTGTGACCTGGCAGCCATCCAAA 1770
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 1771 GATGAGATCGCTTAATTTGGTGGTATTATTAATAAGCGCTTTAAAGATGGAGTGT 1830
QY 681 ProArgAspSerGlyValaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 1831 CCTCAGATTTCAGGAGCAATGCTGGCTTTAAGGTTGTAGGAGGAAGATGACTGAATCA 1890
QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 1891 GGTCCACTTTGTGCATTTATTACCAAGTGAAGAGGAGGTTTGTAGCTGTAGTA 1950
QY 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnIleValaThr 740
Db 1951 CATCTTAGACCAAGTGATGAAGCTTTGGAATGGAATGGAAGCTATTGCAAGGAGCCACA 2010
QY 741 PheGluGluValTyAsnIleLeuGlnSerLysProGluProGlnValGluLeuVal 760
Db 2011 TTTGAGGAGTTTACCAATTATTCTAGAACTCAAGCCTGAACCAAGTTGAGCTTGT 2070
QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2071 GTTTCAGGCCGATGGAGATGCTAGATACCTGATAGCACTCATGCACTGCAACTGGAA 2130
QY 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2131 TCCAGTTCTAGCTCATTTGAATCTCAAAAAATGGATCGTCGCTATATCGGTTACCTCT 2190
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2191 CCCATGAGTCTGGCATGCTGAGGATGTCACACAGATTCTTATCTGGACAGCTTTCAATA 2250
QY 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyValaLysAsp 840
Db 2251 AAATATGTTTGACAGTGGTCCACCAATTAATAGTTTACAAATTTGGAGCAAGGAT 2310
QY 841 LeuProSerArgGluAspGlyArgProArgAsnProTyTrpValLysIleTyPheLeuPro 860
Db 2311 CTCCCTTCCAGGAGATGGGAGGCCAAGGAATCCTTATGTTTAAATTTACTTCTTCCA 2370
QY 861 AspArgSerAspLysAsnLysArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2371 GACAGAAGTGATAAAAAACAGAGAGAGAAACAAAAACAGTCAAGAAACTTTGGAAACCCAAA 2430
QY 881 TrpAsnGlnThrPheIleTyTrpSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2431 TGGAAACCAAGCTTTTCAATTTATCTCTGTTTCCACCAAGAGAAATCCGGGAAGCAATGCTC 2490
QY 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920
Db 2491 GAAATCAACCTTTGGGATCAAGCTCGAGTTCGAGAGGAAGATGAAATTCCTAGAGAG 2550
QY 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTrpLysLeuGln 940
Db 2551 ATTTAAATTTGAATTTGAAACAGCTCTGTTTAGATGATGAACCAACACTGGGTACAACTTCAG 2610
QY 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyTrpLeuProArgArgGln 960
Db 2611 ACCCATGAGTCTCTCATTTGCACTTCCCTTCCATATATGCTTCCACGAGAGAG 2670
QY 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db 2671 CTGCATGGAGAGGCCACACGCGAGTTGCAAGGTGCAAGAGAAATAGTGCACAGCGNA 2730

QY 981 ValSerAspTyTrpAspCysGluAspGlyValGlyValValSerAspTyTrpArgHisAsnGly 1000
Db 2731 GTATCCGACTATGACTGGAGGACGGGTGGGATCGTGTGAGATTACCGACACGATGGC 2790
QY 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
Db 2791 CGTGATCTTCAAAGCTCCACATTATCAGTGCAGACCAAGTCATGTCTCAACCACTTGC 2850
QY 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
Db 2851 TCACCATCAGGCTCTCCTCATCGAGTAGATGTTATAGGAAGGACTAGTGTATGGTCCCT 2910
QY 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 2911 AGTGCTCTCTCTCTCAAGGAATGTGAACAGGGGCTTCGAGGACACAGTGTACTGGC 2970
QY 1061 HisTyTrpAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyTrpSer 1080
Db 2971 CATTAACAATCAATTAGCCGATGGACAGACACCGTGTCTATGGACGACCATTACTTCTCA 3030
QY 1081 AspArgAsp----- 1083
Db 3031 GAGAGACAGCATCTTTTCTCACTCTACCTCTCTCGACACAGGCGACAGCCAGTGAGCAT 3090
QY 1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyTrpHisArgSerArg 1098
Db 3091 CACCACAGGATGGAAGGATTTGAGACGACATAGACAGCCATATCAAGATCCAGA 3150
QY 1099 SerThrGluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArg 1118
Db 3151 TCAACAGAAACAAGGCTCTCTTAGAGCGGACCAACCCGCTCCAGATCTCTGAAAGT 3210
QY 1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
Db 3211 GCTGATACAAACCTCATGAGTGTGATGCTTCAATTAATGATGGAAGATCTGCCCCCTCT 3270
QY 1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
Db 3271 TCACCTGCTTATCGAGTCTCACCTCGCACTGGCTCTGTCCAGACAGCCGCTCAAGT 3330
QY 1159 ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeu 1178
Db 3331 ACTCCGTAAACAGACGAGGGGGCGGCGAGCTTCTCTCAGCTCCCAACCAAGGAGACATG 3390
QY 1179 GluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyLys 1198
Db 3391 GAAAGA----- 3396
QY 1199 GlnValAlaGlySerAspProArgLeuGluGlnAspTyTrpHisSerLysTyTrpArgSerGly 1218
Db 3396 ----- 3396
QY 1219 TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerSerAspSerValSer 1238
Db 3396 ----- 3396
QY 1239 AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyMet 1258
Db 3396 ----- 3396
QY 1259 SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet 1278
Db 3396 ----- 3396
QY 1279 GlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGly 1298
Db 3396 ----- 3396
QY 1299 AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu 1318
Db 3396 ----- 3396
QY 1319 GlyThrSerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGly 1338

QY 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
 DB 229 -----AAG 231
 QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGluGluSerGln 120
 DB 232 CTGCATCAACAAATTTCAATGATTAAGAGCAAGTCAAGAAATGGGAGCAAGTCACAA 291
 QY 121 GlnGlnGlnGlnLysGluAspAlaProThrCysGluVileCysHisLysThrLysPhe 140
 DB 292 CAGCAGCAGAGCAGAGGCGGACCCCGACCTGGCGCATCTGCCACAGCAAAATTT 351
 QY 141 AlaAspGlyCysGluHisAsnCysSerTyrCysGlnThrLysPheCysAlaLysCysGly 160
 DB 352 CGAGATGATCGCGCCATCAATGTTTCAATTCGCAAAACCAAGTCTGTGCTGTGTGCA 411
 QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
 DB 412 GGTGAGTGTCTATTACGCTCAAAACAGGTATGTGGTGTGTAAATTTGTGCCGAAACAA 471
 QY 181 GlnGluIleLeuThrLysSerGluValaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
 DB 472 CAAGAAATCCTCACAATGTCGGCGCGTGTGTTTCAATGATGGGTCTTAATACCGCAG 531
 QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluLysAlaProGlnGlnLysLys 220
 DB 532 CAGCCTGATCAAAAGGCTCTTCAGGCGCTTCAGTGGAGAGCCCTCCAGGAGAGAG 591
 QY 221 AlaLysLeuHisGlnGlnProGlnPheGlnGlnValaProGlyAspLeuSerValProAla 240
 DB 592 GCAAATCTGCATGAGCAGACGAGTTCAGGAGACCCCGGTGACTCATCATCTGCTGCA 651
 QY 241 ValGluLysGluArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
 DB 652 GTTGAGAGAGGTCGAGCTCATGGCTCACAAGCAGGATTCATTAAAGAAATGGCTCAGA 711
 QY 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 DB 712 ATGAAGCACCAGATCCGCAATGATCCCTCCAGACAGAAAGAGTCCGTCACTGTCC 771
 QY 281 ArgAspGlnAsnArgTyrGluGlnSerGluGluArgLysAspTyrSerGlnTyrVal 300
 DB 772 AGGGATCAAAATCGAAGATACCAACCAAGTGAAGAGAGAGAGAAATTAATTCAGATGTT 831
 QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320
 DB 832 CTTTCAGATAGCAGATGCTGATCTCCATCAGATTATGCTGATAGACGATCTCAACGT 891
 QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
 DB 892 GAACCTCAATTTATGAAGACCTGATCATTTAATTTATAGGATTTCAACAGGAGAGGC 951
 QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
 DB 952 CATAGACATTCCAAAGATATATTGTAGACGACGAAGTGTGGAGCAGCAGATGAATAT 1011
 QY 361 GluArgGlnArgArgGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
 DB 1012 GAAGACAAAGAGAGAGAGGAGTACCAGGACGCTACAGAGTGATCCAAATTTGGCC 1071
 QY 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
 DB 1072 CGGTATCCGTTAAGCCCAACCCCTATGAAGCAAAATCGGATCCACGCTGAAGTGTCC 1131
 QY 401 ArgAlaArgHisGluArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
 DB 1132 CGGGCAGCAGCAGCAGAGAGGACAGTGTGTTCTTTGGCAATGCTGAACTAGAGAT 1191
 QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
 DB 1192 TCCAGGATTTCTCTGTATAGGATGGATAGACCATCAAGCAAGATCTGTATCTGAACGT 1251
 QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460

DB 1252 AGAGCTGCAATGGAAACACGAGATCGTATTCAATGGAAAGAACTCGAGAGCTCAGGA 1311
 QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
 DB 1312 CAAAGTCTTATCCACAAGAGCACCACAATCATATAGTCTCTTACCCCTCGCAGAGCCCT 1371
 QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLysLeu 500
 DB 1372 ATACCGCTCGATAGACGAGCTGAGGCTGCCGACTCCCTACGGAACCAACACCCACTTA 1431
 QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
 DB 1432 GATCCACGCTCTGCTGTAAAGGAAACGAGCGAGAAAAATGGAACCAATGTTGAGGAAT 1491
 QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProProProHisLys 540
 DB 1492 GATCTTTGAGTTCAGACCACTGCTGAGTCAGTGGCGGCCCCGCCACCAAGGCCCTCATAA 1551
 QY 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGlnGluLeuAla 560
 DB 1552 TCCAGAAAGAGAGGTTAAATGCGCAGGTTTCACTGACAGCTCAGAGAGAGAGTTGGCA 1611
 QY 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
 DB 1612 TCCACGCTGAGTATACAACTGTGATGCTGAGATTTGAAGCGAGAGCGTAGTGTGAG 1671
 QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlnGlyValLeuSerAspSer 600
 DB 1672 AAAGAGACATGGAG----- 1686
 QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
 DB 1687 -----TACAGCTGGTTGGAG 1701
 QY 621 AspLeuGluTrpSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 640
 DB 1702 CATGGCTCTGGCAT----- 1716
 QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
 DB 1717 -----AGCAGTGGAGCATCCCAATGCTTTGGCACCCCTGTGACCTGGCGCAGCATCAA 1770
 QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuLeuLysArgLeuLysAspGlySerVal 680
 DB 1771 GATGAGATCGCTTAATTTGGTCGTATTTTAAATAAGCGTTTAAAGATGGGAGTGT 1830
 QY 681 ProArgAspSerGlyAlaMetLeuLysValValGlyGlyLysMetThrGluSer 700
 DB 1831 CCTCGAGATTCCAGGAGCAATCTGGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA 1890
 QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
 DB 1891 GGTGACATTTGTGCATTTTATCCAAAGTGAAGGAAAGGAGTTTACGTGATCTAGCA 1950
 QY 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
 DB 1951 CATCTTAGACAGGTCATGAGTCTTGAAATGGAATGGAAGCTTATTCAGAGACCA 2010
 QY 741 PheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVal 760
 DB 2011 TTTGAGGAAGTTTACCAATTTTCTAGATCCCAAGCTGAAACCAACAGTTGAGCTGT 2070
 QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
 DB 2071 GTTTCAGGCGGATTTGGAGATATGCTTAGAATACCTGTAGTACCTCATGCACTGAA 2130
 QY 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
 DB 2131 TCCAGTCTTAGCTCAATTTGAATCTCAAAAAATGGATCGTCCGTCTATATCGGTACCTCT 2190
 QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSer --- 819

Db 2191 CCCATGAGTCTCGCATGCTGAGGGATGTCCACAGTTCTTATCTGGACAGCTTTCAAGC 2250
QY 820 -----1lelylsLeuTrpPhe 824
Db 2251 CAAAGCCTTAGTGAAGAACCAACGCCCTTTTGTCTTAGGGTTCCAGATATAAATATGTTTT 2310
QY 825 AspLysValGlyHisGlnLeuValThrIleLeuGlyAlaLysAspLeuProSerArg 844
Db 2311 GACAGAGTTGGTACACGATTAATAGTTTACAAATTTGGAGCAAGAGATCTCCCTTTCAGG 2370
QY 845 GluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuProAspArgSerAsp 864
Db 2371 GAAGATGGAGGCCCAAGGAATCTTATGTTAAATTTACTTCTTCAGACAGAGTGTAT 2430
QY 865 LysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLysTrpAsnGlnThr 884
Db 2431 AAAAACAAGAGAAGAACAAAACAGTCAAGAAAATTTTGGAAACCAATGGAACCCAGACT 2490
QY 885 PheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluIleThrLeu 904
Db 2491 TTCAATTTATCTCTGTTCCAGGAAGAAATTCGGGAACGAATGCTCGAAATCACCCCTT 2550
QY 905 TrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGluIleLeuIleGlu 924
Db 2551 TGGGATCAAGCTCGAGTTCGAGAGGAAGATGAATTCCTAGGAGAGATTTTAATTGAA 2610
QY 925 LeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGlnThrHisAspVal 944
Db 2611 TTGAAAACAGCTCTGTTAGATGATGAACCAACCACTGGTACAAATTCAGACCCATGATGTC 2670
QY 945 SerSerLeuProLeuProArgProSerProTyrLeuProArgArgGlnLeuHisGlyGlu 964
Db 2671 TCCTCATGCCACTCCCTTCACCTTCTCCATATATGCCAGACAGCTGCATGGAGAG 2730
QY 965 SerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGluValSerAspTyr 984
Db 2731 AGCCCAACACGACAGTTGCCAAGGTCGAGAGATAAGTCACACGGAAGTATCCGACTAT 2790
QY 985 AspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGlyArgAspLeuGln 1004
Db 2791 GACTCGCAGACGGCGTGGAGTCTGCTCAGATTACCACACAGCAGTGGCCGTGATCTTCAA 2850
QY 1005 SerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSerProSerGly 1024
Db 2851 AGCTCCACATTATCAGTGCCAGAACAAATGTCATCAACCAACATTGCTCACCATCAGGC 2910
QY 1025 SerProHisArgValAspValIleGlyArgThrArgSerTrpSerProSerAlaProPro 1044
Db 2911 TCTCTCATCGAGTAGATGTTATAGGAAGGACTAGGTCACTGCTCGCCTAGTCCCTCT 2970
QY 1045 ProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnThr 1064
Db 2971 CCTCAAAGGAATGTGGAAACAGGGCTTCGAGGGACAGTGTACTGCGCCATTACATACA 3030
QY 1065 IleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSerAspArgAsp--- 1083
Db 3031 ATTAGCGGATGGACAGACACCGTGTATGGACGACCATTAATCTCTCAGAGAGAGACAGT 3090
QY 1083 -----1083
Db 3091 CATTTTCTCACTCTACCTCGCTCTCGACACAGGCAGACCATGAGCATCACACAGGGAT 3150
QY 1084 ---ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThrGluGln 1102
Db 3151 GGAAGGGATTGTGAAGCAGCAGATAGACAGCATATCACAGATCCAGATCAACAGAACAA 3210
QY 1103 ArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAspThrAsn 1122
Db 3211 CGGCTCTCTTAGACGGGACCAACACCGCTCCAGATCCTCTGAACGTCGTGTATACAAAC 3270
QY 1123 LeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAlaLeu 1142
Db 3271 CTCATGAGTCTGATGCCCTTCATTAATGACTGGAAAGATCTGCCCTCTCTTACCTGCCCTTA 3330

QY 1143 SerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrProGlyThr 1162
Db 3331 TCAGAGTCTACCCCTCGACTGTCTGTCCAGACAGTCCGTCAGAGTACTCCCGTAACA 3390
QY 1163 GlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArgSerAla 1182
Db 3391 GGACGAAGGGCGCGAGCTTCTCTCAGCTCCCAACAAAGGACACTGGAAGA-----3444
QY 1183 MetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGly 1202
Db 3444 -----3444
QY 1203 SerAspProArgLeuGluAspTyrHisSerLysTyrArgSerGlyTrpAspProHis 1222
Db 3444 -----3444
QY 1223 ArgGlyAlaAspThrValSerThrLysSerSerAspSerValSerAspValSerAla 1242
Db 3444 -----3444
QY 1243 ValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerValGlnSer 1262
Db 3444 -----3444
QY 1263 GluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsnArgGln 1282
Db 3444 -----3444
QY 1283 MetGlyValSerGlyAsnLeuThrLysSerThrSerIleSerGlyAspMetCysSer 1302
Db 3444 -----3444
QY 1303 LeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGly 1322
Db 3444 -----3444
QY 1323 LysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLys 1342
Db 3444 -----3444
QY 1343 SerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLysLeuArgSerThr 1362
Db 3445 -----GGAGGAGTTAAAGCTTACGACGACT 3471
QY 1363 ValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGln 1382
Db 3472 GTCCAGAGAAGCACCGGACCCGGCTTAGCAGTGGAGTGGAGAACTGGATGACAAGACAG 3531
QY 1383 AlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnLeuIle 1402
Db 3532 GCCAGCCGGGAATCCACGGATGGCAGCATGAACAGTTACAGCTCAGAAGGAAATCTGATC 3591
QY 1403 PheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGly 1422
Db 3592 TTTCTCTGGGCTCGCCCTCTGACAGCCAGTTTTCAGGATTTCTTGAATGTCCTGGGC 3651
QY 1423 ProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnVal 1442
Db 3652 CCTGCTCAGCTAGTGGAGCGCCAGACCTCTGGGACTCTCTGCAATGGGTGACATTCAGGTG 3711
QY 1443 GlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArgGlyLeuVal 1462
Db 3712 GCGATGATGATGAAGAAGGACAGCTGGAGGTAGAGATCATCTCGGCTCGAGGCCCTTGTGA 3771
QY 1463 ValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeuAspAsn 1482
Db 3772 GTAAACACCGGTTCCAAAGACACTGCCAGCACCATAATGTCAAGGTGTACCTGTGTAGACAAC 3831
QY 1483 GlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeuGluProLeuTyr 1502
Db 3832 GGAGTCTGCATAGCCAAAGAAACCAAGGTGGCGAAGAGACGCTGGAGCCCTGTAT 3891

891 AAAATGGAAACAATGTTGAGGAATGATCTTTGAGTTTCAGACCAGTCTGAGTCAGTGAGG 950
533 ProProProArgProHisLysSerLysGlyLysMetArgGlnValSerLeu 552
951 CCGCCCCCAAGGCTCATTAATCAAGAAAGGAGTAATGCGCCAGGTTTCACTG 1010
553 SerSerSerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspValGlu 572
1011 AGCAGCTCAGAGGAGGTTGGCATCCAGCCTGAGTATACAAGCTGTGATGACGTGGAG 1070
573 LeuGluSerGluSerValSerGluValGlyValSerGlnLysGlyLysArgLysThrSer 592
1071 ATTGAAGCAGAGCGTAGGTGAGAAAGGAGACATGGAG----- 1109
593 GluGlnGlyValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysArgMetTyr 612
1109 ----- 1109
613 TyrGlyGlyHisSerLeuGluGluAspLeuTyrSerGluProGlnLysLysAspSer 632
1110 -----TACAGCTGGTTGGAGCATGCGTCTTTGGCAT----- 1139
633 GlyValAspThrCysSerSerThrThrLeuAsnGluGluHisSerHisSerAspLysHis 652
1140 -----AGCAGTGAGGCATCCCCAATGCTTTGCAC 1169
653 ProValThrTyrGlnProSerLysAspGlyAspArgLeuLeuGlyArgLysLeuLeuAsn 672
1170 CCGTGACCTGGCAGCATCCAAAGATGGAGATCGCTTAATTGGTGTGATTTTATTAAT 1229
673 LysArgLeuLysAspGlySerValProArgAspSerGlyAlaMetLeuGlyLysVal 692
1230 AAGCGTTTAAAGATGGGAGTGTAACCTCGAGATTTCAGGAGCAATGCTGGGCTTAAAGTT 1289
693 ValGlyLysMetThrGluSerGlyValGluCysAlaPheThrLysValLysLys 712
1290 GTAGGAGAAGATGACTGAATCAGGTGCGACTTTGTGCATTTATTACAAAGTGAAGAAA 1349
713 GlySerLeuAlaAspThrValGlyHisLeuArgProGlyAspGluValLeuGluTyrAsn 732
1350 GGAAGTTTAGCTGATCTGTAGGACATCTTAGACAGGTGATGAAGTCTTGGATGGAT 1409
733 GlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnLysLeuLysSerLys 752
1410 GGAAGGCTATTGCAAGGAGCCACATTTGAGGAGTTTACAACTATTATTCAGAAATCCAAG 1469
753 ProGluProGlnValGluLeuValValSerArgProLysGlyAspLysProArgLysPro 772
1470 CCTGAACCAAGTTGAGCTTGTGTTTCAAGGCCGATGGAGATATGCCCTAGAAATACCT 1529
773 AspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAsp 792
1530 GATAGCACTCATGCACAACTGGAATCCAGTTCTAGCTCATTTGAATCTCAAAAATGGAT 1589
793 ArgProSerLysSerValThrSerProMetSerProGlyMetLeuArgAspValProGln 812
1590 CGTCCGCTATATCGGTTACCTCTCCCATGAGTCTCGCATGCTGAGGAGTGTCCACAG 1649
813 PheLeuSerGlyGlnLeuSer----- 819
1650 TTTCTATCTGACAGCTTTCAAGCCAAAGCCTTAGTAGAAGAACAAAGCCTTTTGTTCCT 1709
820 -----LysLysLeuTyrPheAspLysValGlyHisGlnLeuLeuValThrLysLeu 836
1710 AGGGTTTCAGATAAACTATGTTTGACAAAGTTGGTCCAGTGAATTAATAGTTACAATTTG 1769
837 GlyAlaLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIle 856
1770 GGAGCAAGAGTCTCCCTTCCAGGAGAGATGGAGGCCAAGGAATCTTATGTTAAATTT 1829
857 TyrPheLeuProAspSerAspLysAsnLysArgArgThrLysThrValLysLysThr 876
1830 TACTTCTTCCAGACAGAGTATATAAAACCAAGAGAGAAACAAACCAAGTCAAGAAACT 1889

877 LeuGluProLysTyrAsnGlnThrPheLysTyrSerProValHisArgArgGluPheArg 896
1890 TTGGAACCCAAATGGAAACAGACTTTTCAATTATTTCTCTCTTCCAGGAAGAAATTCGG 1949
897 GluArgMetLeuGluLysLeuThrLeuTyrAspGlnAlaArgValArgGluGluSerGlu 916
1950 GAACGAATGCTCGAATCACCTTTTGGATCAAGCTCGAGTTCGAGGAGAAAGTGA 2009
917 PheLeuGlyLysLeuLeuLeuGluThrAlaLeuLeuAspAspGluProHisTyr 936
2010 TTCTTAGGAGAGATTTTAATTGAATTGGAACAGCTCTGTAGATGATGAACCACTGG 2069
937 TyrLysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeu 956
2070 TACAAACTTCAGACCCATGATGTCTCTCATTTGCCACTCCCTCACCCTTCTCCATATAG 2129
957 ProArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgLys 976
2130 CCACGGAGACAGCTGATGAGAGAGACCCCAACAGCAGGTTGCAAGGTCGAAGAAATA 2189
977 SerAspSerGluValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyr 996
2190 AGTGACAGCGAAGTATCCGACTATGACTGCGAGACGCGTGGAGTCTGTGTGATTTAC 2249
997 ArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSer 1016
2250 CGACAGATGGCGGTGATCTTCAAGCTCCACATTTATCATGTGTCAGAACAGTCAATGTC 2309
1017 SerAsnHisCysSerProSerGlySerProHisArgValAspValLysGlyArgThrArg 1036
2310 TCAACCATTTGCTCACCATCAGGCTCTCTCTCATCGAGTAGATGTTTATAGAGAGACTAG 2369
1037 SerTyrSerProSerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThr 1056
2370 TCAATGTGCTAGTGTCTCTCTCTCAAGGAATGTGAACAGGCGCTTCAGGGACA 2429
1057 ArgAlaThrGlyHisTyrAsnThrLysArgMetAspArgHisArgValMetAspAsp 1076
2430 CGTGTACTGGCCATTAACAATACAAATAGCCGATGTCAGACACACCGTGTCTGACGAC 2489
1077 HisTyrSerSerAspArgAsp----- 1083
2490 CATTTCTTTCAGAGAGAGACAGTCAATTTTCTCACTCTACCTCGCTCTCGACACAGCGAG 2549
1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyr 1094
2550 ACCAGTGAGCATCACCAAGGATGGAGGATTTGTGAAGCAGCAGATAGACAGCCATAT 2609
1095 HisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArg 1114
2610 CACAGATCCAGATCAACAGAACACCGCTCTCTCTAGAGCGGACCAACCCCGCTCCAGA 2669
1115 SerSerGluArgProSerThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArg 1134
2670 TCTCTGAACGCTGTATACAACCTCATGAGGTGATGCTCTCATTAATGACTGGAGA 2729
1135 SerAlaProProSerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThr 1154
2730 TCTGCCCCCTCTTCACTGCTTATCGAGGTCTCACCTCGACTGCTGTGTCCAGACC 2789
1155 SerProSerSerThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProPro 1174
2790 AGTCCGTCAGTACTCCCGTAAACAGGACGAAGGCGCGGAGCTTCTCTAGCTCCCAACA 2849
1175 LysGlyThrLeuGluArg----- 1180
2850 AAGGGACACTGGAAAGAAATGATCACAGAAGACATGCACTCCCAAGGAAGAGAACTCA 2909
1181 SerAlaMetAspLysGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
2910 GGTGCTATGATATAGAGGAGAAATCGCCAAATGAAATTTAACAAATTAACACAGGTA 2969

1201 AlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpAsp 1220
Db 2970 GCGGATCAGATCCAGACTGAGACAGATACCATTCGAGATGCTCAGATGGGAT 3029
1221 ProHisArgGlyAlaAspThrValSerThrLysSerAspSerAspValSerAspVal 1240
Db 3030 CCACATAGAGGGGCGAGATCTGTTCTACTAAATCCTCGACAGATGATGAAGTATGTA 3089
1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerThrMetSerVal 1260
Db 3090 TCTCGGTTTCAAGACTAGTAGTCTCTCTGTTTTCAGCAGACACAGCTACATGTCGTC 3149
1261 GlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsn 1280
Db 3150 CAATCAGACGGCGGAGAGAACAGAAATCAGTGTCTTACGTCCTCAAAATGCAAGC 3209
1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMet 1300
Db 3210 AGACAGATGGGTGTGTCGGGGAAGACATGGCCAAAAGCAACAGCATCAGTCGAGCATG 3269
1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValAlaLeuGlyThr 1320
Db 3270 TGTCTCAGTGGAGAAATGATGCGACCGACCTGCGACCTGCGGCGCCCTGGGCGACC 3329
1321 SerGlyLysLysArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
Db 3330 AGTGGCAGACGGCGATCTAGCATTTGGGGCCAAATAGTAGCTATTGTTGCTCTCA 3389
1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArg 1360
Db 3390 CGGAAAAGTCCAGTGCCTCCCACTCAGCCAAACCGAAGGAGGAGTAAAGCTACGC 3449
1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
Db 3450 AGCACTGTCCAGAGAAGCAGCGAGACCGGGCTAGCAGTGGAGATGAGGAACCTGATCA 3509
1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerGluGlyAsn 1400
Db 3510 AGACAGCGCCAGCGGAAATCCAGGATGGCAGATGAACAGTTACAGCTCAGAGGAAT 3569
1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
Db 3570 CTGATCTTCTGGGGTCCGCTGGCTCTGACAGCCAGTTCCAGCATTTCTTGTATGT 3629
1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
Db 3630 CTGGGCGCTGCTCAGCTAGTGGAGCGCCAGACCTGGCGACTCTCTCAATGGGTGACATT 3689
1441 GlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaGly 1460
Db 3690 CAGGTGGGATGATGATGAAGAGGACAGCTGAGGTAGAGATCATCCGGGCTCGAGGC 3749
1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu 1480
Db 3750 CTGTAGTAAACAGGTTCCAGACACTGCCAGCACCATATATGTCAGGTGTACCTGTTA 3809
1481 AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeuGluPro 1500
Db 3810 GACACGAGTCTGATAGCCAAAAGAAACCAAGTGGCGGAGAGAGAGCGTGGAGCGCC 3869
1501 LeuTyrGlnGlnLeuSerPheGluGluSerProGlnGlyValGluGlnIleIle 1520
Db 3870 CTGTATCAGCAGCTGTATCTTTCGAAGAGAGAGCCCGAGGAGGTGTGTACAGATCAT 3929
1521 ValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
Db 3930 GTCTGGGAGATTTATGGCGCATGATCATCAATCTTTATGGAGTGGCCGACATCTC 3989
1541 LeuAspGluLeuGluSerAsnMetValIleGlyTrpPheLysLeuPheProSer 1560
Db 3990 TTAGATGAATCGAATCACTCAACATGATGTAATGGATGTTCAAACTCTTCCCTCTCC 4049
1561 SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerLeuGlu 1580

Db 4050 TCCCTAGTAGATCAACCTTGGCACTCTGACAGAGAGCTTCCCAATGCTCTCGAA 4109
Qy 1581 SerSerThrGlyProSerTyrSerArgSer 1590
Db 4110 AGTTCTACTGACCTTCTTACTCTGTTCA 4139
RESULT 15
AF199328 4851 bp mRNA linear ROD 04-JUL-2000
LOCUS Rattus norvegicus RIM2-4B (Rim2-4B) mRNA, complete cds.
DEFINITION AF199328
ACCESSION AF199328
VERSION AF199328.1 GI:8925869
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 4851)
Wang, Y., Sugita, S. and Sudhof, T.C.
The RIM/NIM family of neuronal C2 domain proteins. Interactions
with Rab3 and a new class of Src homology 3 domain proteins
J. Biol. Chem. 275 (26), 20033-20044 (2000)
JOURNAL MEDLINE 20347919
PUBMED 10748113
REFERENCE 2 (bases 1 to 4851)
Wang, Y. and Sudhof, T.C.
Direct Submission
TITLE Submitted (27-OCT-1999) Center for Basic Neuroscience, The
University of Texas Southwestern Medical Center, 6000 Harry Hines
Blvd., Dallas, TX 75235-9111, USA
FEATURES
Location/Qualifiers
1..4851
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..4851
/gene="Rim2-4B"
52..3930
/gene="Rim2-4B"
/codon_start=1
/product="RIM2-4B"
/protein_id="AAF81650.1"
/db_xref="GI:8925870"
/translation="MSAPLGRGRPAPTPAASOPPPPEFMDLSHLTBEERKIIQAVM
DRQKEBEKQSVLKHQFQFMYKQVKGESQOQKQGDAPTCGICHTKPAD
CGHNCYCTQKFCARCGRSLRKNVWVNCNCRKQOEILTKSAGWYNSGNTPO
QPDOKALASERAPQEKAKLHBTQFQPGDSSVPAVERGAHGLTRQDSIKNG
SGMKQIADMPSDRKRSPSRDONRQYDQSEEREYSQYPSSTWPSRSDYADR
RSQEPQFYERPDHLYRDSNRGRHSKTYIVDDVEDSEDEYRQREREYQAIR
SDPNLARYPKFPYEQMRTHAEVSRAPHERHSDVSLANLEDSRISLMDRPS
RQSRYSERRAENQSYSMERTEAQCSYPTQTTNHSPTTPRSPTPLDRPELR
ADSLRKHLPDSSAVRTKREKMTLRNDLSLSDQSESVPPPPRPHSKKGGKMR
QVLSLSSSEELASTPEYTSDDVEISESVGKEGMEYSLEHSAHSSSESPMSLHP
VTWQSKDGLRIGILLNKLKDGSVPRDSQMLGLKVGGEMTESGELCAPIYKVK
KGLADTVHLRPGDEVLEWNGLLQGFATFSEVINILASRPEPVELVPSPIGMP
RIPSTHAQLSSSSSSSQMDRPSITFSPMSGLMDRVPQFLSGSLKLMFDKY
GHQIVTLTGALDFSRDGRPNPYKVIYFLDRSDKNKRTKTKVTKLEPNQTF
IYSPVRRERFRERLEITLMDQARVREESFLGELILELALDDPHWYKLTQTH
VSSLPLHPSYPMRQLHGSPTLRQRKRIISDSVSDYDCEDGVGVSDYRDRGR
DLQSTLSVPEQVMSNCSPSGSPHYDVTIGRTSNPSFPVPPPNVQGLRGTAT
GHYNTLSMDRHRVMDHYSSRDSHFULPSRHRQTSVQSEETGLAVENWMTQASRE
PSSTVTRGRQLPQFPTLGRGKKLRSTVQSEETGLAVENWMTQASRE
STDSGMSYSGENLIPFVRLASDSQFSLDGLGPAQLVGRQTLATPAMQIQVM
MDKKQLSVIIRAGLVVKGSKTLPAVYKVLDDNGVCIKKTKVARKTLRPLLY
QQLLSPESPOKVLQIIVMDYGBMDHKSFWGVAQIILLDELSLNNWLGWFKLPPPS
SLLDPLAPLTRASQSSLESSTGSPSISRS"
BASE COUNT 1453 a 1110 c 1209 g 1079 t
ORIGIN
Alignment Scores: 4.11e-187 Length: 4851
Pred. No.:

Score:	6212.00	Matches:	1226
Percent Similarity:	78.43%	Conservative:	21
Best Local Similarity:	77.11%	Mismatches:	45
Query Match:	74.98%	Indels:	298
DB:	10	Gaps:	5

US-09-617-099B-1 (1-1590) x AF199328 (1-4851)	
QY 1 MetSerAlaProLeuGlyProArgGlyProAlaProThrProAlaLaserGlnPro 20	
DB 52 ATGTGGCTCCACTCGGGCCCGGGCGCGCCGCCCTCCACCCCGCGGCTCTCAGCCT 111	
QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysLeu 40	
DB 112 CCTCCGAGCCCGAGATCGCGGACCTCAGCCACTCAGGAGAGGAGGAGAAATCATC 171	
QY 41 LeuAlaValMetAspArgGlnLysLeuGluGluGluGluGluGluGluLysVal 60	
DB 172 CAGGCTGTCTATGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228	
QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyLeuThrGluLeu 80	
DB 228 ----- 228	
QY 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGlnLysGluProGlnThrLys 100	
DB 229 -----AAG 231	
QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120	
DB 232 CTGCATCAACAATTTGAATGTATPAGAGCAAGTCAAGAAATCGGAGAGGATCAAA 291	
QY 121 GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyLeuCysHisLysThrLysPhe 140	
DB 292 CAGCAGCAAGAGCAGAGGCGCGCCCGACCTCGCGCATCTGCCCAACAAGCAAAATTT 351	
QY 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160	
DB 352 GCAGATGGATCGCGCCCACTAATGTTTATATTTGCCAAACCAAGTTCTGTCTCGTGTGGA 411	
QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180	
DB 412 GGTGAGTGTATTACGCTCAAAACAGGTATGTGGTGTGTATTTGTGCCGGAACAA 471	
QY 181 GlnGluLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200	
DB 472 CAAGAATCCTCACTAAGTCGGCGCGTGTGTATCAATAGTGGTCTAATACACCCGAG 531	
QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLysLys 220	
DB 532 CAGCCTGATCAAAAGGCTCTTCGAGGGCTTCGAGTGGAGAGGCCCTCAGGAGAGAAG 591	
QY 221 AlaLysLeuHisGlnGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240	
DB 592 GCAAACTGCTATGAGACACCGAGTTCAGGAGACCCCGGTGACTCATAGTACTTGCA 651	
QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrLysAsnGlySerGly 260	
DB 652 GTTCAGAGAGTTCGAGCTCATGGCTCACAAGACAGGATTCTATTAGAATGGCTCAGGA 711	
QY 261 ValLysHisGlnLysAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280	
DB 712 ATGAAGCACCCAGATCGCCAGTACATGCTTCAGACAGAAAGAGAGTCCGTCAGTGTCC 771	
QY 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300	
DB 772 AGGGATCAAAATCGAAGATACGACCAAGTGAAGAAAGAGAGGAGATATTTCAGTATGTT 831	
QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320	
DB 832 CCTTCAGATAGACACAAATGCTATGATCTCATCAGATTATGCTGATAGACGATCTCAAGT 891	
QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgGly 340	

DB 892 GAACTCTAATTTATGAGAGACTGATCATTTAAATTTATAGGATTTCTAACAGGAGGC 951	
QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360	
DB 952 CATAGACATTTCCAAAGAGTATATTGTAGACGACGAGAGTGTGAGAGCGAGATGAAT 1011	
QY 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380	
DB 1012 GAAGACAAGGAGAGAGAGAGGAGTACCAAGGACGCTCAGAGTGTATCCAATTTGGCC 1071	
QY 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400	
DB 1072 CCGTATCCGTTAAGCCACACCTTATGAAGCAATTCGGATCCACCTGAGTGTCTC 1131	
QY 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420	
DB 1132 CGGCGACGACACGAGAGAGGACAGTGTATTTCTTGGCAATGTCTGAAGTATGAAGAT 1191	
QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440	
DB 1192 TCCAGGATTTCTCTGCTAAGGATGGATAGACCATCAAGGCAAGATCTGTATCTGAACGT 1251	
QY 441 ArgAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460	
DB 1252 AGAGCTGCAATGGAACCCAGCGATCTGTATTCATGGAAGAACTCGAAGCTCAGGA 1311	
QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480	
DB 1312 CAAAGTTCTTATCCCAAGGACCACAATCATAGTCTCTCTACCTCCGAGGAGCCCT 1371	
QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500	
DB 1372 ATACCCCTCGATAGACGAGCTGAGGCGTGGCGACTTCCCTACGGAACACACCACTTA 1431	
QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520	
DB 1432 GATCCAGCTCTCTCTGTAAAGGAAACGAGCGAGAAAATATGGAACAATGTTGAGAAAT 1491	
QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540	
DB 1492 GATTCITTTGAGTTTCAGACCATCTGAGTCAGTGAGCGCGCCCTCACCAAGGCTCATAAA 1551	
QY 541 SerLysLysGlyLysMetArgGlnValSerLeuSerSerSerSerGluGluLeuAla 560	
DB 1552 TCCAAGAAAGAGAGTAAATGCCAGGTTTCTACAGCAGCTCAGAGGAGGAGTGGCA 1611	
QY 561 SerThrProGluTyrThrSerCysAspValGluLeuGluSerGluSerValSerGlu 580	
DB 1612 TCCAGCGCTGAGTATACAGCTGTGATGAGTGGAGATTGAAAGCGAGAGCGTAGGTGAG 1671	
QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600	
DB 1672 AAAGGACATCGAG----- 1686	
QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620	
DB 1687 -----TACAGCTGTGGAG 1701	
QY 621 AspLeuGluTrpSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640	
DB 1702 CATGCGTCTTGGCAT----- 1716	
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660	
DB 1717 -----AGCAGTGGAGCATCCCAATGCTCTTTCACCTCTGACCTGGCAGCATCCAAA 1770	
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680	
DB 1771 GATGGAGATCGCTTAATTTGCTGCTATTTTATTAATAAGCGTTTAAAGATGGAGTGA 1830	
QY 681 ProArgAspSerGlyAlaMetLeuGlyLeuValValGlyLysMetThrGluSer 700	

Qy 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
Db 3418 CTGGCCCTGCTCAGCTAGTGGGAGCGCCAGACCTCGCGACTCTCGAATGGGTGACATT 3477
Qy 1441 GlnValGlyMetMetAspLysLysGlyClnLeuGluValGluIleIleArgAlaArgGly 1460
Db 3478 CAGGTGGGATGATGGATAGAGGGACAGCTGGAGGTAGAGATCATCCGGGCTCGAGGC 3537
Qy 1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu 1480
Db 3538 CTTCGTAGTAAACACAGGTTCACAGACACTGCCAGCACCATATGTCAAGGTGTACCTGTTA 3597
Qy 1481 AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeuGluPro 1500
Db 3598 GACAAACGGAGTCTGCATAGCCCAAAAGAAACCAAGGTGGCGAGAGAGACCTGGAGCCC 3657
Qy 1501 LeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleIle 1520
Db 3658 CTGTATCAGCAGCTGTTATCTTTCGAGAGAGACCCCGAGGGGAGGTGTACAGATCATT 3717
Qy 1521 ValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
Db 3718 GTCTGGGAGATTATGGCCGCATGGATCACAAATCCTTTATGGAGTGGCCCGACATATC 3777
Qy 1541 LeuAspGluLeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPheProSer 1560
Db 3778 TTAGATGAACCTGGAACCTATCCAACTATGGTAAATGGATGGTTCAAACCTCTCCCTCTCC 3837
Qy 1561 SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGlu 1580
Db 3838 TCCCTAGTAGATCCAACTTGGCACCTCTGACAGAGAGGCTTCCCAATCGTCTCTGGAA 3897
Qy 1581 SerSerThrGlyProSerTyrSerArgSer 1590
Db 3898 AGTCTACTGGACCTTCTTACTCTCGTTCA 3927

Search completed: November 16, 2003, 13:14:57
Job time : 13295 secs